

Sami Savio

**MAGNEETTIKUVIIN SOVELLETUN
TEKSTUURIANALYYSIN TULOKSIIN
VAIKUTTAVIEN TEKIJÖIDEN
ARVIOINTIA**

Lääketieteen ja terveysteknologian tiedekunta
Lisensiaatintutkimus
Lokakuu 2019

TIIVISTELMÄ

Sami Savio: Magneettikuviin sovelletun tekstuurianalyysin tuloksiin vaikuttavien tekijöiden arviointia

Lisensiaatintutkimus, 55 sivua, 33 liitesivua

Tampereen yliopisto

Biolääketieteen tekniikan tohtorihjelma

Lokakuu 2019

Magneettikuvantamisen yleisiä käyttökohteita ovat potilaan päähän, sisäelimiin ja tukirankaan kohdistuvat tutkimukset. Tekstuurianalyysin avulla magneettikuvien yksityiskohtia on mahdollista analysoida numeerisesti, mistä voi olla suurta hyötyä potilaiden hoidossa.

Tässä työssä tarkasteltiin lääketieteellisen tutkimuksen ja klinisen lääketieteen apuna käytettävän, magneettikuviin sovelletun tekstuurianalyysin teknisten edellytysten suhdetta tekstuurianalyysin tarkkuuteen ja luotettavuuteen. Tutkimuksen kohteena olivat erityisesti magneettikuvien leikepaksuus (Osatyö I) ja niiden signaali-kohinasuhde (Osatyö II) sekä tekstuurianalyysin yhdistäminen diffuusiotensorikuvaukseen (Osatyö III).

Aineistona käytettiin multippeliskleroosia sairastavien potilaiden aivoista ja terveiden vapaaehtoisten tutkimuskohteiden reiden alueesta otettuja magneettikuvia. Tutkimustulosten yleistettävyyttä pyrittiin parantamaan ristiinvalidoinnin avulla erillisiä opetus- ja testiaineistoja käyttäen.

Tutkimustulosten perusteella tekstuurianalyysi osoittautui suhteellisen vakaaksi menetelmäksi. Pienillä muutoksilla leikepaksuudessa tai signaali-kohinasuhteessa oli vain vähäinen vaikutus tekstuurianalyysin tarkkuuteen. Diffuusiotensorikuvauksen käyttö yhdessä tekstuurianalyysin kanssa ei parantanut multippeliskleroosin tautityyppiluokittelun tarkkuutta, vaikka paljastikin taudin etenemiseen viittaavia eroja vasemman ja oikean aivopuoliskon välillä.

Tehdyt havainnot viittaavat siihen, että tekstuurianalyysillä on edellytyksiä yleistyä ja antaa luotettavia tuloksia kliinisessä hoitotyössä. Tekstuurianalyysin vakiintuminen osaksi päivittäistä hoitorutiinia edellyttää silti yhä jatkotutkimusta sen tarkkuuteen ja tulosten toistettavuuteen vaikuttavista tekijöistä kaikissa tekstuurianalyysin vaiheissa. Erityistä huomiota tulee kiinnittää näiden tekijöiden vaikutuksen kasautumiseen varsinkin silloin, kun ne ovat riippuvuussuhteessa toisiinsa.

Avainsanat: Tekstuurianalyysi, magneettikuvaus, leikepaksuus, signaali-kohinasuhde, diffuusiotensorikuvaus

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

ABSTRACT

Sami Savio: Estimating factors that influence the results of the texture analysis of magnetic resonance images

Licentiate thesis, 55 pages, 33 Appendix pages

Tampere University

Doctoral Programme in Biomedical Sciences and Engineering

October 2019

Magnetic resonance imaging is commonly used for examining patients' heads, viscera and spinal columns. Texture analysis enables the quantification of the details in magnetic resonance images, which may improve treatment.

This study investigated the technical requirements for the texture analysis of the magnetic resonance images used in medical research and clinical medicine with respect to the accuracy and the reliability of the analysis. The objectives of the study were the slice thickness (Study I) and signal-to-noise ratio (Study II) in the magnetic resonance images along with the combination of texture analysis and diffusion tensor imaging (Study III).

The study data included magnetic resonance images of the brain area of multiple sclerosis patients and the thigh area of healthy subjects. Cross-validation with separate training and test data sets was used to enhance the generalizability of the results.

The results indicate that texture analysis is a relatively stable method. Small changes in the slice thickness or signal-to-noise ratio produced only minor changes in the accuracy of the analysis. The combination of the diffusion tensor imaging and texture analysis methods did not improve the accuracy of the separation of multiple sclerosis subtypes even though it revealed some hemispheric differences in disease progression.

The findings of this study suggest that clinical texture analysis may provide reliable results in daily practice and become more common. However, before texture analysis becomes a part of daily clinical routines, additional studies on the factors that affect its accuracy throughout the analysis as well as the repeatability of the results are necessary. The cumulative effects of these factors must be evaluated, especially when they are cross-dependent.

Keywords: Texture analysis, magnetic resonance imaging, slice thickness, signal-to-noise ratio, diffusion tensor imaging

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

ALKUSANAT

Tähän työhön liittyvä tutkimus on toteutettu pääosin Tampereen yliopistollisessa sairaalassa. Kiitän lämpimästi kaikkia tutkimustyöhön osallistuneita henkilöitä ja varsinkin professori Hannu Eskolaa, joka on toiminut myös lisensiaatintutkimukseni vastuuhjaajana. Olen iloinen saatuani osallistua vuosien saatossa moneen kiinnostavaan ja tärkeään tutkimusprojektiin.

Osoitan lopuksi erityiskiitokset perheelleni ja vanhemmilleni kaikesta vuosien myötä saamastani tuesta ja kannustuksesta sekä opinnoissani että niiden ulkopuolella.

Ylöjärvellä 28.9.2019

Sami Savio

SISÄLLYSLUETTELO

TIIVISTELMÄ	I
ABSTRACT	II
ALKUSANAT	III
SISÄLLYSLUETTELO	IV
LYHENTEET JA MERKINNÄT	VI
LUETTELO ALKUPERÄISJULKAISUISTA	IX
1. JOHDANTO	1
2. TEKSTUURIANALYYSIN PERUSTEITA	3
2.1 Yleisesti käytettyjä tekstuurimenetelmiä	4
2.2 Tekstuuriparametrien valinta	8
2.3 Tekstuurianalyysin yhteydessä käytettäviä luokittelualgoritmeja	9
2.4 Teksturiin perustuva segmentointi	10
3. MAGNEETTIKUVIEN TEKSTUURIANALYYSI	11
3.1 Magneettikuvien tekstuurianalyysin sovelluskohteita	11
3.2 Magneettikuvien kolmiulotteinen tekstuurianalyysi	13
3.3 Tekstuurianalyysia täydentäviä menetelmiä	14
4. MAGNEETTIKUVIEN TEKSTUURIANALYYSIN TARKKUUTEEN	
VAIKUTTAVIA TEKIJÖITÄ	16
4.1 Signaali-kohinasuhde	16
4.2 Leikepaksuus ja resoluutio	18
4.3 Magneettikuvantamisen sekvenssien ja parametrien valinta	19
4.4 Kontrasti	21
4.5 Magneettikuvan intensiteettivääritykset	21
4.6 Tarkasteltavan ROI-alueen valinta	22
5. AINEISTOT JA MENETELMÄT	24
5.1 Leikepaksuuden vaikutus tekstuurianalyysin tarkkuuteen	25
5.2 Signaali-kohinasuhteen vaikutus tekstuurianalyysin tarkkuuteen	26
5.3 Tekstuurianalyysin ja diffuusiotesorikuvauksen yhdistäminen	28
6. TUTKIMUSTULOKSET	31
6.1 Leikepaksuus	31
6.2 Signaali-kohinasuhde	32
6.3 Tekstuurianalyysin ja diffuusiotesorikuvauksen yhdistäminen	36
7. POHDINTAA	38
7.1 Leikepaksuus	38

7.2 Signaali-kohinasuhde	39
7.3 Tekstuurianalyysin ja diffuusiotensorikuvauksen yhdistäminen.....	41
7.4 Muiden tekijöiden vaikutus tekstuurianalyysin tarkkuuteen	42
8.JOHTOPÄÄTÖKSET	45
LÄHDELUETTELO.....	46
ALKUPERÄISJULKAISUT	55

LYHENTEET JA MERKINNÄT

ADC	Näennäinen diffuusiovakio (<i>apparent diffusion coefficient</i>)
CT	Tietokonetomografia (<i>computed tomography</i>)
DTI	Diffuusiotensorikuvaus (<i>diffusion tensor imaging</i>)
FA	Fraktionaalinen anisotropia (<i>fractional anisotropy</i>)
FLASH	<i>Fast low angle shot</i> -sekvenssi
GLCM	Yhteisesiintymämatriisi (<i>grey-level co-occurrence matrix</i>)
GLRLM	Jaksonpituusmatriisi (<i>grey-level run-length-matrix</i>)
GE	Kenttäkaikusekvenssi (<i>gradient echo</i>)
IR-GE	Käänteispalautumiskenttäkaikusekvenssi (<i>inversion recovery gradient echo</i>)
IR-SE	Käänteispalautumisspinkaikusekvenssi (<i>inversion recovery spin echo</i>)
k-avaruus	Magneettikuvainformaation taajuus- ja vaihemuodossa sisältävä mittausavaruus
k-NN	k:n lähimmän naapurin menetelmä (<i>k nearest neighbors method</i>)
LBP	Paikallinen binäärikuvio (<i>local binary pattern</i>)
LDA	Lineaarinen erotteluanalyysi (<i>linear discriminant analysis</i>)
MPR	Monisuuntareformaatti (<i>multiplanar reconstruction</i>)
MRI	Magneettikuvaus (<i>magnetic resonance imaging</i>)
MS	Multippeliskleroosi (<i>multiple sclerosis</i>)
MSi	MS-plakin päälle vapaasti piirretty ROI-alue (<i>MS irregular</i>)
MSr	MS-plakin päälle piirretty vakiokokoinen ROI-alue (<i>MS regular</i>)
m.	Lihaskuvaus (<i>muscle</i>)
NAWM	Normaalilta näyttävä valkea aine (<i>normal appearing white matter</i>)
NDA	Epälineaarinen erotteluanalyysi (<i>non-linear discriminant analysis</i>)
NWM	Normaali valkea aine (<i>normal white matter</i>)
PPMS	Ensisijaisesti etenevä multippeliskleroosi (<i>primary progressive multiple sclerosis</i>)
RF	Radiotaajuus (<i>radio frequency</i>)
ROI	Tarkastelualue (<i>region of interest</i>)
RRMS	Aaltomainen multippeliskleroosi (<i>relapsing-remitting multiple sclerosis</i>)
SE	Spinkaikusekvenssi (<i>spin echo</i>)
SNR	Signaali-kohinasuhde (<i>signal-to-noise ratio</i>)
SPMS	Toissijaisesti etenevä multippeliskleroosi (<i>secondary progressive multiple sclerosis</i>)
SVM	Tukivektorikone (<i>support vector machine</i>)
TA	Tekstuurianalyysi (<i>texture analysis</i>)
TE	Kaikuaika (<i>time to echo</i>)
TI	Inversioaika (<i>time of inversion</i>)
TR	Toisto-aika (<i>time of repetition</i>)
T1	Magnetisaation pitkittäinen relaksaatioaika
T1C	T1-painotteinen, kuvan kontrastia (<i>contrast</i>) parantavaa varjoainetta hyödyntävä sekvenssi
T2	Magnetisaation poikittainen relaksaatioaika
T2*	Kenttäkaikusekvensseille tyypillinen magnetisaation poikittainen relaksaatioaika
VOI	Tarkastelutilavuus (<i>volume of interest</i>)
WM	Valkea aine (<i>white matter</i>)
2D	Kaksiulotteinen (<i>two-dimensional</i>)
3D	Kolmiulotteinen (<i>three-dimensional</i>)
3 σ	Kolmen keskihajonnan (<i>three-sigma</i>) etäisyys

$ABSV(i,j)$	Absoluuttinen gradientti
b	Yhteisesiintymämatriisin määrittelyvektorin pituus
C	Jaksonpituusmatriisista johdettu vakio
D	Luokkien välinen hajonta
F	Fisherin lukuarvo
e_s	Autoregressiivisen mallin tuottama virhe-estimaattori
f_s	Kuvapisteen s intensiteetti
N	Kohinateho
N_g	Kuvan harmaasävyjen lukumäärä
N_r	Harmaasävyiltään yhtenäisten jaksojen lukumäärä
N_s	Kuvapisteen s ympäristö
$p(i)$	Todennäköisyys tapahtumalle tai harmaasävyille i
$p(i,j)$	Tapahtumien tai harmaasävyjen i ja j yhteisesiintymän todennäköisyys
$p_x(i)$	Harmaasävyyn i täsmälleen j :ssä peräkkäisessä pikselissä esiintymiskertojen lukumäärä
$p_y(j)$	
$p_{x+y}(k)$	
$p_{x-y}(k)$	
$q(i,j)$	
S	Signaaliteho
V	Luokkien sisäinen hajonta
$x(i,j)$	Tason kuvapiste
β	Yhteisesiintymä- tai jaksonpituusmatriisin määrittelyvektorin suunta- kulma
θ_r	Autoregressiivisen mallin tuottama parametrivektori
μ	Keskiarvo
σ	Keskihajonta
σ^2	Varianssi
μ_3	Jakauman vinous (<i>skewness</i>)
μ_4	Jakauman huipukkuus (<i>kurtosis</i>)
μ_x, μ_y, μ_{x-y}	Yhteisesiintymämatriisista johdettuja keskiarvoja
σ_x, σ_y	Yhteisesiintymämatriisista johdettuja keskihajontoja
$GrMean$	Gradienttiparametreja
$GrVariance$	
$GrSkewness$	
$GrKurtosis$	
$GrNonZeros$	
$Fraction$	Jaksonpituusmatriisiparametreja
$GlevNonUni$	
$LngREmph$	
$RLNonUni$	
$ShrtREmph$	

<i>AngScMom</i>	Yhteisesiintymämatriisiparametreja
<i>Contrast</i>	
<i>Correlat</i>	
<i>DifEntrp</i>	
<i>DifVarnc</i>	
<i>Entropy</i>	
<i>InvDfMom</i>	
<i>SumAverg</i>	
<i>SumEntrp</i>	
<i>SumOfSqs</i>	
<i>SumVarnc</i>	

LUETTELO ALKUPERÄISJULKAISUISTA

- Osatyö I: [I] Savio SJ, Harrison LCV, Luukkaala T, Heinonen T, Dastidar P, Soimakallio S, Eskola HJ. Effect of slice thickness on brain magnetic resonance image texture analysis. Biomedical Engineering Online 2010; 9:60.
- Osatyö II: [II] Savio S, Harrison L, Ryymin P, Dastidar P, Soimakallio S, Eskola H. Estimation of sufficient signal to noise ratio for texture analysis of magnetic resonance images. Proceedings of SPIE 7962, Medical Imaging 2011: Image Processing, 79622C.
- Osatyö III: [III] Savio S, Hakulinen U, Ryymin P, Hagman S, Dastidar P, Soimakallio S, Elovaara I, Eskola H. Hemispheric asymmetry measured by texture analysis and diffusion tensor imaging in two multiple sclerosis subtypes. Acta Radiologica 2015; 56: 844-851.

1. JOHDANTO

Magneettikuvaus (magnetic resonance imaging, MRI) on nopeasti yleistynyt lääketieteellisen kuvantamisen muoto. Sen yleisiä käyttökohteita ovat muun muassa potilaan pää-
hän, sisäelimiin ja tukirankaan kohdistuvat tutkimukset.

Magneettikuvausten yleistymisestä huolimatta edes lääketieteellisten kuvantamistutkimusten analysointiin kouluttautunut radiologi ei havaitse läheskään kaikkia kuvan yksityiskohtia, kuten monimutkaisia muotoja, rakenteita tai tekstuureja. Tekstuurianalyysin (texture analysis, TA) avulla lääketieteellisten kuvien yksityiskohtia on mahdollista analysoida numeerisesti, mistä voi olla merkittävää hyötyä potilaiden hoidossa ja hoidon seurannassa [1].

Tämän työn tavoitteena on tarkastella lääketieteellisen tutkimuksen ja kliinisen lääketieteen apuna käytettävän, magneettikuviin sovelletun tekstuurianalyysin teknisten edellytysten suhdetta sen tarkkuuteen ja luotettavuuteen. Tutkimuksen kohteena ovat erityisesti magneettikuvantamisessa käytettävä leikepaksuus (Osatyö I, [I]) ja magneettikuvien signaali-kohinasuhde (signal-to-noise ratio, SNR) (Osatyö II, [II]) sekä tekstuurianalyysin yhdistäminen diffuusiotensorikuvaukseen (diffusion tensor imaging, DTI) (Osatyö III, [III]).

Luvussa 2 kuvataan tekstuurianalyysin perusteita, ja Luvussa 3 sen magneettikuviin liittyviä sovelluksia. Luvussa 4 esitellään magneettikuvien tekstuurianalyysin tarkkuuteen, luotettavuuteen ja toistettavuuteen vaikuttavia tekijöitä. Niitä ovat edellä mainittujen seikkojen lisäksi muun muassa kuvantamisessa käytettyjen parametrien valinta, magneettikentän epähomogeenisuus, tarkasteltavan ROI-alueen (region of interest) muoto ja koko sekä kuvan resoluutio ja kontrasti.

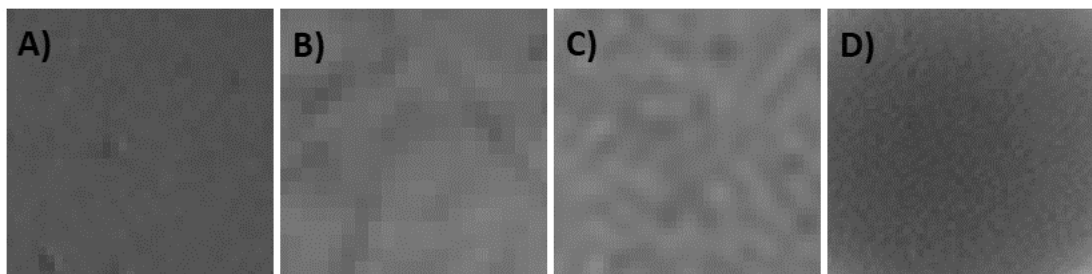
Tutkimusmenetelmät ja -aineistot esitellään Luvussa 5, joka perustuu kolmeen aiemmin julkaistuun, magneettikuvien tekstuurianalyysin alalle sijoittuvaan tieteelliseen artikkeliin [I-III]. Osatyössä I [I] (Luku 5.1) on tarkasteltu magneettikuvien leikepaksuuden ja Osatyössä II [II] (Luku 5.2) niiden signaali-kohinasuhteen vaikutusta lääketieteellisen tekstuurianalyysin tarkkuuteen. Osatyössä III [III] (Luku 5.3) on tutkittu, voidaanko tekstuurianalyysin tarkkuutta parantaa soveltamalla sitä yhdessä diffuusiotensorikuvauksen

kanssa. Tutkimusaineistoina käytettiin multippeliskleroosia (multiple sclerosis, MS) sairastavien potilaiden aivoista sekä terveiden vapaaehtoisten tutkimuskohteiden reiden alueesta otettuja magneettikuvia.

Tutkimustulokset esitellään Luvussa 6, ja niiden merkitystä analysoidaan yksityiskohtaisemmin Luvussa 7. Luku 8 sisältää työn yhteenvedon ja johtopäätökset.

2. TEKSTUURIANALYYSIN PERUSTEITA

Tekstuurilla ei ole tarkkaa ja yksiselitteistä määritelmää, mutta sen voidaan katsoa muodostuvan esimerkiksi erilaisten kuvioiden tai ryhmitelmien muodostamista säännöllisistä joukoista [2]. Luonnollisten kohteiden kuvat sisältävät lukuisia erilaisia tekstuurikuvioita. Mahdollisesti monimutkaisetkin intensiteettikuviot voivat olla muun muassa jatkuvia tai epäjatkuvia, säännöllisiä tai epäsäännöllisiä sekä hieno- tai karkeajakoisia. Kuvassa 1 on esitetty lihaskudoksen, rasvakudoksen ja aivojen valkean aineen sekä lääketieteellisten kuvantamislaitteiden laadunvalvonnassa käytetyn polyamidifantomien tekstuurit.



Kuva 1. A) Lihaskudoksen, B) rasvakudoksen, C) aivojen valkean aineen ja D) polyamidifantomien tekstuurit.

Silmämääräisen arvioinnin lisäksi tekstuuria voidaan tulkita numeerisesti tekstuurianalyysin avulla. Kuvien numeerisesta tulkinnasta on monien muiden sovellusalojen ohella hyötyä myös lääketieteellisessä käytössä. Koska tekstuurianalyysi tuottaa lääketieteellisistä kuvista kvantitatiivista informaatiota, se voi auttaa potilaskuvien tulkinnassa ja lääketieteellisten hoitojen seurannassa [3-4].

Tekstuurianalyysi kattaa laajan joukon erilaisia tekniikoita ja sovelluksia [5-6]. Tekstuurianalyysimenetelmän valinta riippuu yleensä sovelluskohteesta. Larrozan ym. mukaan lääketieteellisten kuvien tekstuurianalyysissä on kuitenkin tyypillisesti kuusi erillistä, toisiaan seuraavaa vaihetta: 1) kuvantaminen, 2) analysoitavien ROI-alueiden valinta, 3) ROI-alueiden esikäsittely, 4) tekstuuriparametrien laskenta, 5) analyysissä käytettävien tekstuuriparametrien valinta ja 6) kudosten algoritminen luokittelu tekstuuriparametrien avulla [1]. Kukin näistä vaiheista vaikuttaa osaltaan tekstuurianalyysin luotettavuuteen ja tarkkuuteen.

Tekstuurianalyysissä voidaan hyödyntää esimerkiksi kuvan harmaasävyeroja mittaavia tilastollisia parametreja. Niitä ovat muun muassa histogrammi-, gradientti-, yhteisesiintymämatriisi- (grey-level co-occurrence matrix, GLCM) ja jaksonpituusmatriisiparametrit

(grey-level run-length-matrix, GLRLM) [2]. Edellä mainittuja menetelmiä on käytetty laajasti tässä työssä, ja niitä kuvaillaan tarkemmin Luvussa 2.1.

Tekstuurimenetelmät voidaan ryhmitellä rakenteellisiin, malliperustaisiin, tilastollisiin ja erilaisiin muunnoksiin perustuviin lähestymistapoihin. Tekstuurianalyysi voi yksinkertaisempien menetelmien lisäksi perustua esimerkiksi fraktaaleihin, regressioanalyysiin tai wavelet-muunnokseen. Laskennallisesti hienostuneimpia menetelmiä vaivaa usein monimutkaisuus. [7] Wavelet-muunnoksen puutteena pidetään sen epäintuitiivisuutta [8].

Tekstuurimenetelmiä voidaan tarvittaessa yhdistellä. Eräs esimerkki rakenteellisten ja tilastollisten tekstuurimenetelmien yhdistelmästä on LBP-menetelmä (local binary pattern) [9]. Koneoppimisen arvioidaan laajentavan entisestään tekstuurianalyysin tarjoamia mahdollisuuksia [10]. Hájekin ym. toimittamassa teoksessa on esitetty yksityiskohmainen kuvaus useista erilaisista tekstuurianalyysimenetelmistä sovelluskohteineen [2].

Lerski ym. ovat laatineet kokoelman suosituksia tekstuurianalyysin käytöstä [11]. He nostivat samassa yhteydessä esiin kolme olennaista huomiota. Lerskin ym. mukaan tärkeintä tekstuurianalyysia sovellettaessa on ymmärtää syvällisesti se, mitä etsitään ja lasketaan. He myös painottivat tulosten toistettavuuden merkitystä. Kolmanneksi Lerski ym. varoittivat tutkijoita niin sanotusta liiallisten ulottuvuuksien kirouksesta (curse of dimensionality) ja luokittelualgoritmin ylisovituksen (overfitting) vaaroista. Mainittuja riskejä voidaan yrittää välttää vähentämällä luokittelussa käytettyjen tekstuuriparametrien määrää sekä soveltamalla esimerkiksi erillisiä opetus- ja testiaineistoja hyödyntävää ristiinvalidointia. Näin on menetelty myös tässä työssä.

2.1 Yleisesti käytettyjä tekstuurimenetelmiä

Ensimmäisen kertaluvun tilastollisiin menetelmiin kuuluvilla histogrammiparametreilla voidaan muun muassa mitata kuvan tai sen osan harmaasävyjen keskiarvoa μ , varianssia σ^2 , harmaasävyjakauman vinoutta (skewness) μ_3 tai huipukkuutta (kurtosis) μ_4 [2]:

$$\mu = \sum_{i=1}^{N_g} ip(i) \quad (1)$$

$$\sigma^2 = \sum_{i=1}^{N_g} (i - \mu)^2 p(i) \quad (2)$$

$$\mu_3 = \sigma^{-3} \sum_{i=1}^{N_g} (i - \mu)^3 p(i) \quad (3)$$

$$\mu_4 = \sigma^{-4} \sum_{i=1}^{N_g} (i - \mu)^4 p(i) - 3 \quad (4)$$

Edellä $p(i)$ on harmaasävyn i esiintymistodennäköisyys ja N_g käytössä olevien harmaasävyjen lukumäärä.

Toisen ja korkeamman kertaluvun tilastolliset menetelmät edellyttävät myös tietoa kuvan harmaasävyjen sijainnista. Esimerkiksi gradienttiin perustuvat tekstuuriparametrit toimivat intensiteettimuutosten mittana. Merkitään kuvapisteen $x(i,j)$ kaksiulotteista 3x3-naapurustoa seuraavasti [2]:

G	H	I
L	$x(i,j)$	N
Q	R	S

Tämän jälkeen voidaan määritellä absoluuttinen gradientti $ABSV(i,j)$ [2]:

$$ABSV(i,j) = \sqrt{(R - H)^2 + (N - L)^2} \quad (5)$$

sekä siitä johdetut parametriarvot M pikseliä sisältävässä ROI-alueessa [2]:

$$GrMean = \frac{1}{M} \sum_{i,j \in ROI} ABSV(i,j) \quad (6)$$

$$GrVariance = \frac{1}{M} \sum_{i,j \in ROI} (ABSV(i,j) - GrMean)^2 \quad (7)$$

$$GrSkewness = \frac{1}{GrVariance^{\frac{3}{2}} M} \sum_{i,j \in ROI} (ABSV(i,j) - GrMean)^3 \quad (8)$$

$$GrKurtosis = \frac{1}{GrVariance^2 M} \sum_{i,j \in ROI} (ABSV(i,j) - GrMean)^4 - 3 \quad (9)$$

$$GrNonZeros = [ABSV(i,j) > 0]\% \quad (10)$$

Jaksonpituusmatriisiparametreja voidaan hyödyntää kuvan yhtenäisten harmaasävyjaksojen analyysissä [2]. Olkoon jatkossa $q(i,j)$ niiden kertojen määrä, kun harmaasävy i esiintyy täsmälleen j :ssä peräkkäisessä pikselissä suunnassa β (0, 45, 90 tai 135 astetta). Olkoon lisäksi N_g edelleen mahdollisten harmaasävyjen lukumäärä ja N_r harmaasävyiltään yhtenäisten jaksojen lukumäärä sekä

$$C = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} q(i,j) \quad (11)$$

Jaksonpituusmatriisiparametrit voidaan nyt määritellä seuraavasti (menetelmä on yleistettävissä kolmeen ulottuvuuteen) [2]:

$$Fraction = C / \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j q(i,j) \quad (12)$$

$$GLevNonUni = \frac{1}{C} \sum_{i=1}^{N_g} \left(\sum_{j=1}^{N_r} q(i,j) \right)^2 \quad (13)$$

$$LngREmph = \frac{1}{C} \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j^2 q(i,j) \quad (14)$$

$$RLNonUni = \frac{1}{C} \sum_{j=1}^{N_r} \left(\sum_{i=1}^{N_g} q(i,j) \right)^2 \quad (15)$$

$$ShrtREmph = \frac{1}{C} \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{q(i,j)}{j^2} \quad (16)$$

Yhteisesiintymämatriisiparametrit perustuvat harmaasävyparien samanaikaisen esiintymisen todennäköisyysjakaumaan. Yhteisesiintymämatriisiparametreja on useita erilaisia [2]:

$$AngScMom = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)^2 \quad (17)$$

$$Contrast = \sum_{n=0}^{N_g-1} n^2 \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)_{|i-j|=n} \quad (18)$$

$$Correlat = \frac{1}{\sigma_x \sigma_y} \left(\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} ijp(i,j) - \mu_x \mu_y \right) \quad (19)$$

$$DifEntrp = - \sum_{i=1}^{N_g} p_{x-y}(i) \log(p_{x-y}(i)) \quad (20)$$

$$DifVarnc = \sum_{i=0}^{N_g-1} (i - \mu_{x-y})^2 p_{x-y}(i) \quad (21)$$

$$Entropy = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \log(p(i,j)) \quad (22)$$

$$InvDfMom = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i,j)}{1 + (i-j)^2} \quad (23)$$

$$SumAverg = \sum_{i=1}^{2N_g} ip_{x+y}(i) \quad (24)$$

$$SumEntrp = - \sum_{i=1}^{2N_g} p_{x+y}(i) \log(p_{x+y}(i)) \quad (25)$$

$$SumOfSqs = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)^2 p(i,j) \quad (26)$$

$$SumVarnc = \sum_{i=1}^{2N_g} (i - SumAverg)^2 p_{x+y}(i) \quad (27)$$

Edellä N_g on käytössä olevien harmaasävyjen lukumäärä,

$p(i,j)$ todennäköisyys intensiteettien i ja j yhteisesiintymälle suunnassa β ja etäisyydellä b ,

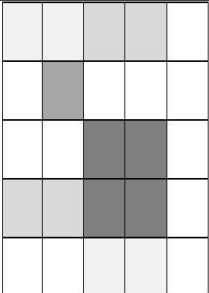
$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)_{i+j=k}, k = 2, 3, \dots, 2N_g \quad (28)$$

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)_{|i-j|=k}, k = 0, 1, \dots, N_g - 1 \quad (29)$$

$p_x(i)$ ja $p_y(j)$ reunajakaumat (marginal distributions) sekä

$\mu_x, \mu_y, \mu_{x-y}, \sigma_x$ ja σ_y vastaavat keskiarvot ja keskihajonnat.

Yhteisesiintymämatriisiparametrit voidaan laskea 0, 45, 90 ja 135 asteen suunnissa sekä esimerkiksi 1-5 pikselin etäisyyksillä. Lisäksi yhteisesiintymämatriisiparametrit voidaan jaksonpituusmatriisiparametrien tavoin yleistää kolmeen ulottuvuuteen. Kuvassa 2 on esitetty 5*5 pikselin kokoinen esimerkkikuva pikseli-intensiteetteineen, siitä lasketut GLCM- ($b = 1, \beta = 0^\circ$) ja GLRLM-matriisit ($\beta = 0^\circ$) sekä neljän näistä matriiseista johdetun tekstuuriparametrin arvot.

	<table><tr><td>1</td><td>1</td><td>2</td><td>2</td><td>0</td></tr><tr><td>0</td><td>3</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>4</td><td>4</td><td>0</td></tr><tr><td>2</td><td>2</td><td>4</td><td>4</td><td>0</td></tr><tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td></tr></table>	1	1	2	2	0	0	3	0	0	0	0	0	4	4	0	2	2	4	4	0	0	0	1	1	0	<table><tr><td>4</td><td>1</td><td>0</td><td>1</td><td>1</td></tr><tr><td>1</td><td>2</td><td>1</td><td>0</td><td>0</td></tr><tr><td>1</td><td>0</td><td>2</td><td>0</td><td>1</td></tr><tr><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>2</td><td>0</td><td>0</td><td>0</td><td>2</td></tr></table>	4	1	0	1	1	1	2	1	0	0	1	0	2	0	1	1	0	0	0	0	2	0	0	0	2	<table><tr><td>5</td><td>2</td><td>1</td><td>0</td><td>0</td></tr><tr><td>0</td><td>2</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>2</td><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>2</td><td>0</td><td>0</td><td>0</td></tr></table>	5	2	1	0	0	0	2	0	0	0	0	2	0	0	0	1	0	0	0	0	0	2	0	0	0
1	1	2	2	0																																																																										
0	3	0	0	0																																																																										
0	0	4	4	0																																																																										
2	2	4	4	0																																																																										
0	0	1	1	0																																																																										
4	1	0	1	1																																																																										
1	2	1	0	0																																																																										
1	0	2	0	1																																																																										
1	0	0	0	0																																																																										
2	0	0	0	2																																																																										
5	2	1	0	0																																																																										
0	2	0	0	0																																																																										
0	2	0	0	0																																																																										
1	0	0	0	0																																																																										
0	2	0	0	0																																																																										
Esimerkkikuva	Pikseli-intensiteetit	GLCM ($b = 1, \beta = 0^\circ$)	GLRLM ($\beta = 0^\circ$)																																																																											
LngREmph = 3,1333 / Fraction = 0,6 / AngScMom = 0,1 / Contrast = 3,85																																																																														

Kuva 2. Esimerkkikuva, sen pikseli-intensiteetit, GLCM- ja GLRLM-matriisit sekä neljän tekstuuriparametrin arvot.

Kuten aiemmin todettiin, on tekstuurianalyysissä mahdollista hyödyntää monia muitakin menetelmiä. Esimerkiksi yksinkertainen autoregressiivinen malli on määriteltävissä seuraavasti [2]:

$$f_s = \sum_{r \in N_s} \theta_r f_r + e_s \quad (30)$$

Edellä f_s on kuvapisteen s intensiteetti, N_s kuvapisteen s ympäristö, θ_r mallin parametrivektori ympäristössä N_s ja e_s mallin estimointivirhe. Autoregressiivisen mallin tuottamia tekstuuriparametreja ovat vektorin θ_r alkiot ja estimaattori e_s .

2.2 Tekstuuriparametrien valinta

Jirák ym. ovat pitäneet tekstuuriparametrien valintaa magneettikuvien tekstuurianalyysin kannalta ratkaisevana vaiheena [12]. Heidän mukaansa on lähes mahdotonta arvioida etukäteen, mitkä parametrit ovat biologisten kudosten luokittelun kannalta kaikkein käytökelpoisimpia. Gyuon ja Elisseef ovat todenneet, että vaikka jokin parametri olisikin luokittelun kannalta yksinään täysin hyödytön, se saattaa olla hyödyllinen toisiin paramet-

reihin yhdistettynä [13]. Kuten edellä mainittiin, samanaikaisesti käytössä olevien parametrien määrää tulee kuitenkin rajoittaa [11]. Esimerkiksi Larroza ym. ovat kehottaneet laskemaan mahdollisimman monta tekstuuriparametriarvoa ja valitsemaan käyttötarkoituksen kannalta tehokkaimmat parametrit jatkoanalyysiin [1].

Depeursingen ym. mukaan tekstuuriparametrien tehokas valinta edellyttää etukäteistietoa käsiteltävästä aineistosta. Parametrien sisäisen skaalan ja suuntien tulee yhtyä analysoitavan datan rakenteen kanssa. Tarkasteluetaisyydet ovat usein pieniä, korkeintaan kolmen pikselin suuruisia. [14] On kuitenkin huomioitava, että tekstuuriparametrien etäisyyksien muuntelu mahdollistaa tekstuurien havainnoinnin eri skaaloilla [15].

Fisherin menetelmä on tekstuurianalyysin yhteydessä yleisesti käytetty parametrinvalintakriteeri. Siinä vertaillaan tarkastelun kohteena olevien luokkien välistä ulkoista hajontaa D niiden sisäiseen hajontaan V ja valitaan ne parametrit, joilla ulkoisen ja sisäisen hajonnan välinen suhdeluku F on suurin [2]:

$$F = \frac{D}{V} \quad (31)$$

Fisherin menetelmän laskennallinen ja tilastollinen skaalautuvuus tekee siitä Guyonin ja Elisseeffin mukaan suositeltavan muihin parametrinvalintakriteereihin verrattuna [13]. Erityisen huomionarvoista on, että Fisherin menetelmä on suhteellisen immuuni datan ylisovitusriskille [13][16].

2.3 Tekstuurianalyysin yhteydessä käytettäviä luokittelualgoritmeja

Eräs tekstuurianalyysin tärkeimmistä käyttökohteista on tarkasteltavien kohteiden luokittelu niiden sisältämän tekstuurin perusteella. Lineaarinen erotteluanalyysi (linear discriminant analysis, LDA) kuuluu yleisimpiin tekstuurianalyysin yhteydessä käytettäviin luokittelualgoritmeihin.

Suhteellisen yleisiä menetelmiä ovat myös epälineaarinen erotteluanalyysi (non-linear discriminant analysis, NDA) sekä diskreetti k :n lähimmän naapurin menetelmä (k nearest neighbors, k -NN). Esimerkiksi muotojen yhteneväisyyttä arvioitaessa k -NN-menetelmää on pidetty toimivana luokittelualgoritmina [17]. Tekstuurianalyysin yhteydessä on hyödynnetty useita muitakin luokittelumenetelmiä, muun muassa neuroverkkoa ja tukivektorikonetta (support vector machine, SVM) [6][18].

2.4 Tekstuuriin perustuva segmentointi

Tekstuuriin perustuva segmentointi ja tekstuurianalyysi ovat lääketieteellisiin kuviin sovellettujen tekstuurimenetelmien kaksi yleisintä käyttökohdetta. Rathore ym. ovat laatineet katsauksen tekstuuriin perustuviin maksakudoksen analyysi- ja segmentointitekniikoihin [19]. Heidän mukaansa tekstuurimenetelmän valinta riippuu käyttötarkoituksen ohella kuvantamismodaliteetista.

Tekstuuriin perustuvaa segmentointia voidaan käyttää esimerkiksi jakamaan lääketieteellisen kuvan sisältö eri kudoksia sisältäviin alueisiin. Segmentointi voi olla joko ohjattua (supervised) tai ohjaamatonta (unsupervised) [3]. Campadelli ym. ovat kehittäneet täysautomatisoidun algoritmin maksan tekstuuriperusteiseen segmentointiin [20].

Lääketieteellisten kuvien segmentointi tähtää anatomisten rakenteiden tutkimiseen, tarkastelun kohteena olevien alueiden identifiointiin, kudosten tilavuuden mittaamiseen tai hoitosuunnitelmien teon avustamiseen [21]. Tekstuuriin perustuva segmentointi voidaan jakaa kahteen modulaariseen prosessiin, jotka ovat piirteiden mittaaminen ja piirteiltään yhtenäisten alueiden segmentointi [22].

Pham ym. ovat analysoineet lääketieteellisten kuvien segmentointimenetelmien hyviä ja huonoja puolia. Niin sanottu osittaistilavuusvaikutus (partial volume effect) hankaloittaa erityisesti matalaresoluutioisten kuvien segmentointia. Lisäksi magneettikuville tyypillinen intensiteettien epähomogeenisuusvääristymä hankaloittaa varsinkin niiden automaattista segmentointia. [23]

Segmentointitarkkuutta voidaan yrittää kohentaa eri menetelmiä ja lähestymistapoja yhdistelemällä. Muneeswaranin ym. mukaan tekstuuriin perustuvan segmentoinnin tarkkuutta on mahdollista parantaa yhdistämällä kuvan avaruudellisia ja taajuusominaisuuksia [24]. Kekre ja Gharge ovat ehdottaneet kuvan intensiteettien korvaamista niiden todennäköisyysjakaumalla [25]. Tesar ym. ovat tutkineet kaksi- ja kolmiulotteisia segmentointialgoritmeja. He ovat ehdottaneet yhteisesiintymämatriisien painottamista kaksiulotteisista ROI- tai kolmiulotteisista VOI-alueista (volume of interest) mitatun etäisyyden funktiona [26].

Kruggel ym. segmentoivat 116 ikääntyneen henkilön aivoista otettuja magneettikuvia, jotka sisälsivät patologisia löydöksiä [27]. He käyttivät yhteisesiintymämatriiseja diffuusin valkean aineen hyperintensiivisten alueiden ja periventrikulaaristen vaurioiden segmentointiin sekä hyödynsivät odotusarvon maksimointialgoritmia arvioidakseen yhteisesiintymämatriisien parametrijakaumaa. Kruggelin ym. mukaan vastaavaa segmentointialgoritmia voitaisiin soveltaa esimerkiksi MS-potilaiden aivokuviin.

3. MAGNEETTIKUVIEN TEKSTUURIANALYYSI

Magneettikuvaus on yleistynyt nopeasti lääketieteellisessä käytössä [28]. Eräs merkittävä syy magneettikuvien runsaasti lisääntyneeseen hyödyntämiseen on se, että monesta muusta radiologiassa käytetystä menetelmästä poiketen magneettikuvantamisessa ei synny ionisoivaa säteilyä [29-30].

Muiden lääketieteellisten kuvien tavoin magneettikuviin sisältyy useita monimutkaisia muotoja ja rakenteita, tekstuureja. Tietokoneella suoritettavan tekstuurianalyysin avulla kuvien yksityiskohtia voidaan analysoida numeerisesti ja mahdollisesti tuottaa ihmisille näkymätöntä lisätietoa. [1]

Magneettikuvien tekstuurianalyysi tähtää tutkittavien kudosten tai alueiden karakterisointiin ja luokitteluun. Tekstuurianalyysillä on lääketieteessä lukuisia sovelluskohteita, esimerkiksi aivovauriot ja -sairaudet, kuten kasvaimet [31], epilepsia [32-33], Alzheimerin tauti [34-35], Parkinsonin tauti [36] ja MS-tauti [37], syöpätaudit [38-42], päähän kohdistuneet iskut [43] sekä sisäelinten [44], lihasten [45-48] ja tukirangan [49] toiminta ja häiriöt.

3.1 Magneettikuvien tekstuurianalyysin sovelluskohteita

Aivokasvainten karakterisointi on neurologisiin magneettikuviin sovelletun tekstuurianalyysin varhainen sovellusalue [50-51]. Kassnerin ja Thornhillin mukaan tekstuurianalyysi on erityisen hyödyllinen aivovaurioiden karakterisointimenetelmä [6]. Kruggel ym. ovat ehdottaneet aivokammioiden ympärillä olevien valkean aineen harventumien ja signaalinvoimistumina ilmenevien vaurioiden luokittelua niiden tekstuurin ja anatomisen sijainnin perusteella [27].

Multippeliskleroosi eli MS-tauti on krooninen keskushermoston sairaus [52]. Sekä perinnölliset että ympäristötekijät vaikuttavat MS-taudin etenemiseen [53], ja se luokitellaan kliinisten löydösten perusteella eri muotoihin [54]. Yleisin MS-taudin tyyppi on aaltomainen MS (relapsing-remitting MS, RRMS). Siinä ajoittaisia ja äkillisiä taudin pahenemiskauskoja seuraavat osittaiset tai täysimittaiset toipumisjaksot.

Ellei RRMS-taudin hoitoon kiinnitetä ajoissa riittävää huomiota, muuntuu se usein vaikeammaksi, toissijaisesti eteneväksi SPMS-taudiksi (secondary progressive MS) [55]. Eräs

harvinainen mutta vaikea MS-taudin muoto on ensisijaisesti etenevä MS (primary progressive MS, PPMS), jossa tauti etenee parantumattomasti sen puhkeamisesta lähtien [56]. Minagar ym. ovat kehottaneet kiinnittämään erityistä huomiota MS-taudin varhaisvaiheen muutoksiin harmaassa aineessa ja varsinkin thalamuksessa, koska ne osallistuvat suureen määrään neurologisia toimintoja [57]. Muun muassa Zhangin ym. [4] ja Harrisonin ym. [37] mukaan magneettikuvien tekstuurianalyysistä voidaan kehittää lääketieteellinen menetelmä MS-taudin diagnosointiin ja taudin etenemisen seuraamiseen. Tähän mahdollisuuteen ovat viitanneet myös esimerkiksi Zhang ym. [8].

Bento ym. hyödynsivät lukuisia erilaisia luokittelualgoritmeja aivojen valkean aineen hyperintensiteettiä tutkiessaan [58]. Heidän käyttämiinsä menetelmiin kuuluivat lineaarisen erotteluanalyysin lisäksi muun muassa tukivektorikoneet ja 1-NN-algoritmi sovellettuina yhteisesiintymämatriisi-, jaksonpituusmatriisi- ja wavelet-tekstuuriparametreihin. Jokainen menetelmä pystyi erottelemaan normaalin valkean aineen hyperintensiivisistä MS-plakeista yli 98 % tarkkuudella. Tulehduksista ja hapenpuutteesta aiheutuneiden vaurioiden erottelu osoittautui vaikeammaksi ja onnistui luokittelualgoritmista riippuen noin 85-88 % tarkkuudella. Kussakin luokittelutehtävässä tosin käytettiin vähintään 30 tekstuuriparametria, mitä on pidettävä verrattain korkeana ja datan ylisovittamisen riskiä merkittävästi kasvattavana määränä.

Tekstuurianalyysia voidaan hyödyntää muidenkin kehonosien, esimerkiksi lihasten tutkimisessa. Muun muassa Lerski ym. ovat arvioineet tekstuurianalyysin voivan olla käyttökelpoinen lihaskudosten luokittelumenetelmä [11]. Herlidou ym. ovat raportoineet tekstuurianalyysin pystyvän erottelemaan terveet lihakset rappeutumisesta kärsivistä lihaksista 70 % sensitiivisyydellä ja 86 % spesifisyydellä [45]. Heidän tutkimuksessaan radiologit eivät saavuttaneet läheskään vastaavaa luokittelutarkkuutta. Radiologien tekemien analyysien sensitiivisyys oli 56 % ja spesifisyys 71 %.

Nketiah ym. hyödynsivät tekstuurianalyysia tutkiessaan fyysisen harjoittelun vaikutuksia lihaskudoksiin [48]. Heidän tarkastelemaisina lihaksina olivat iso pakaralihas (m. gluteus maximus), keskimäinen pakaralihas (m. gluteus medius), lonkankoukistaja (m. iliopsoas) ja sisempi peittäjälihas (m. obturator internus). Nketiah ym. havaitsivat urheilijoiden ja kontrolliryhmän välillä useita tilastollisesti melkein merkitseviä ($p < 0,05$) tekstuurieroja. Heidän tekemänsä tutkimuksen perusteella tekstuurianalyysin avulla voidaan havaita esimerkiksi tietyn tyyppisten harjoitusmenetelmien vaikutukset lihaskudosten rakenteeseen. Kestävyysjuoksijoiden ja kontrolliryhmän välillä oli eroja kaikissa neljässä tutkitussa lihasryhmässä. Sitä vastoin voimanostajien ja kontrolliryhmän lihaskudoksista ei löydetty tilastollisesti merkitseviä tekstuurieroja.

3.2 Magneettikuvien kolmiulotteinen tekstuurianalyysi

Tekstuurianalyysi on sovellettavissa sekä kaksi- (2D) että kolmiulotteiseen (3D) kuvaan. Näistä lähestymistavoista kaksiulotteinen tekstuurianalyysi on käytännön syistä selvästi yleisempi magneettikuvien analysointimenetelmä. Useiden tutkimustulosten nojalla tyypillinen ero kaksi- ja kolmiulotteisen menetelmän tarkkuudessa on kuitenkin muutama prosenttiyksikkö kolmiulotteisen menetelmän hyväksi.

Wagner ym. testasivat kolmea toisistaan poikkeavaa menettelytapaa verratessaan 2D- ja 3D-tekstuurimenetelmien tarkkuutta rinnan alueen magneettikuvalöydösten luokittelussa [59]. Ensin he hyödynsivät vain yhtä kaksiulotteista magneettikuvaleikettä potilasta kohti. Seuraavaksi he analysoivat jokaisen potilaan kaikki käyttökelpoiset magneettikuvaleikkeet ja keskiarvoistivat saadut arvot tekstuuriparametrikohteisesti. Kolmannessa menettelytavassa datasta laskettiin myös kolmiulotteisten tekstuuriparametrien arvot.

Wagner ym. totesivat kolmiulotteisen tekstuurianalyysin kaksiulotteista menetelmää keskimäärin tarkemmaksi, joskin kaksiulotteinen lähestymistapa tuotti joissakin tapauksissa parempia tuloksia. Erityisesti kaksiulotteisten tekstuuriparametrien keskiarvoistaminen johti usein vähintään yhtä hyvään luokittelutulokseen kuin kolmiulotteinen tekstuurianalyysi. Tutkijat päättelivät, että suurin osa rintamagneettikuvien tekstuurianalyysin kannalta olennaisesta informaatiosta sisältyi kaksiulotteiseen leiketasoon. [59]

Georgiadis ym. ovat osoittaneet 3D-tekstuurianalyysin olevan 2D-analyysia tarkempi kallon sisäisten kasvainten ja etäpesäkkeiden luokittelumenetelmä [60]. Heidän tutkimuksessaan luokittelutarkkuus koheni 90 %:sta 94 %:iin, kun luokittelussa hyödynnettiin kolmiulotteista tekstuurimenetelmää kaksiulotteisen sijaan. Mahmoud-Ghoneim ym. puolestaan tutkivat aivojen valkeaa ainetta, turvotusalueita, solukuolema-alueita ja aivokasvaimia kaksi- ja kolmiulotteisilla tekstuurianalyysimenetelmillä. He havaitsivat, että kolmiulotteinen tekstuurianalyysi oli selvästi kaksiulotteista menetelmää tarkempi lineaarista erotteluanalyysia hyödynnettäessä. Mahmoud-Ghoneim ym. arvioivat löydöksensä johtuvan ainakin osittain siitä, että kolmiulotteisessa menetelmässä tekstuuriparametreja oli mahdollista laskea suuremmalta tarkastelualueelta. [61]

Chen ym. ovat huomauttaneet, että pyörähdysinvarianssi kolmiulotteisten yhteisesiintymämatrisiparametrien suhteen voidaan saavuttaa laskemalla tekstuuriparametrikohteisesti kaikkien käytössä olevien suuntien parametrikeskiarvo. Heidän tutkimuksessaan rintojen magneettikuviin sovelletun tekstuurianalyysin luokittelutarkkuus oli huomattavasti parempi kolmessa kuin kahdessa ulottuvuudessa. Tarkkuus koheni koko tarkastelualueen tekstuuria hyödynnettäessä. [62]

Nketiahin vertaillessa neljää reiden alueen lihasta oli kaksiulotteisen tekstuurimenetelmän keskimääräinen tarkkuus 86 % ja kolmiulotteisen menetelmän 91 % [63]. Nketiahin tarkastelemien kaksiulotteisten tekstuuriparametrien hajontakuviot olivat suurelta osin päällekkäisiä. Kolmiulotteisessa tekstuurianalyysissä lihasryhmät sitä vastoin erottuivat toisistaan huomattavasti selkeämmin.

3.3 Tekstuurianalyysia täydentäviä menetelmiä

Tekstuurianalyysin tarkkuutta voidaan yrittää parantaa käyttämällä sitä yhdessä muiden, tekstuurista riippumattomien menetelmien kanssa. Eräs magneettikuvien tekstuurianalyysia mahdollisesti täydentävä menetelmä on diffuusiotesorikuvaus [64]. Se on erityinen magneettikuvantamisen muoto, jolla mitataan kudosten molekyyli-diffuusion voimakkuutta sekä suuntautuneisuutta eli anisotropiaa.

Diffuusion anisotropiaa mittaava FA-arvo (fractional anisotropy) on dimensioton ja saa välillä 0 ja 1 olevia arvoja. Täysin isotrooppisen diffuusion, esimerkiksi satunnaisen lämpöliikkeen FA-arvo on nolla ja täysin suuntautuneen diffuusion FA-arvo 1. Diffuusion voimakkuutta mitataan ADC-arvolla (apparent diffusion coefficient), jonka dimensio on mm^2/s . Diffuusion voimakkuus ja suunta voivat muuttua esimerkiksi onnettomuuden tai vamman seurauksena. Diffuusiotesorikuvauksen tuottamia FA- ja ADC-karttoja voidaan hyödyntää esimerkiksi aivovamma- ja MS-potilaiden hoidon suunnittelussa [65-66].

Eliat ym. yhdistivät dynaamisen varjoainetehostetun magneettikuvauksen useiden tekstuurianalyysimenetelmien kanssa luokitellakseen aivokasvaimia T1-painotteisista magneettikuvista [67]. Heidän tutkimuksessaan yhteisesiintymämatriisiparametrien käyttö johti jaksonpituusmatriisiparametrien hyödyntämistä selvästi parempaan tulokseen. Yhteisesiintymämatriisiparametrien ja dynaamisen varjoainetehostetun magneettikuvauksen yhteiskäyttö paransi olennaisesti luokittelun sensitiivisyyttä, mutta ei sen spesifisyyttä. Eliatin ym. mukaan spesifisyyttä voitaisiin kenties parantaa hyödyntämällä T2-painotteisten kuvien tekstuuria, DTI-parametreja tai kolmiulotteista tekstuurianalyysia.

Myös kineettisiä tai morfologisia piirteitä on mahdollista hyödyntää tekstuurianalyysin yhteydessä [62]. Esimerkiksi histologisen datan ja tekstuurin korrelaatio voi olla pieni [68]. Toisaalta de Certaines ym. ovat muistuttaneet histologisten kuvien ja magneettikuvien suurista resoluutioeroista, jotka tekevät niiden yhdistämisestä vaikeaa [69].

Zacharaki ym. käyttivät tukivektorikoneita aivokasvainten luokittelussa [18]. Tekstuuriparametrien lisäksi he hyödynsivät analyysissaan muun muassa kasvainten muotoon liittyviä parametreja. Etäpesäkkeiden ja keskushermoston kasvainten luokittelutarkkuus oli

heidän algoritmillaan 85 %. Matalan ja korkean erilaistumisasteen kasvainten luokittelutarkkuus oli 88 %. Zacharaki ym. ovat ehdottaneet magneettikuvauksen yhdistämistä MR-spektroskopian, mikroskooppikuvauksen tai histopatologisten tutkimusten kanssa aivokasvainten luokittelutarkkuuden parantamiseksi.

4. MAGNEETTIKUVIEN TEKSTUURIANALYYSIN TARKKUUTEEN VAIKUTTAVIA TEKIJÖITÄ

Lääketieteellisten kuvien laadulla on niiden kvantitatiivisen analysoinnin onnistumisen kannalta suuri merkitys. Kuviin kohdistuvat laatuvaatimukset korostuvat myös tekstuurianalyysin lääketieteellisissä sovelluksissa. Puigin ja Gargian mukaan hahmontunnistus-algoritmit ovat riippuvaisia useista käytetyn datan ominaisuuksista [70]. Magneettikuvan datasisältöön vaikuttavat kuvauskohteen lisäksi muun muassa kuvantamissekvenssin ja siihen liittyvien parametrien valinta, kuvaresoluutio ja leikepaksuus sekä kohina ja muut kuvantamistilanteeseen liittyvät häiriölähteet.

Pusey ym. ovat luetelleet useita tekijöitä, jotka synnyttävät vääristymiä (artefacts) magneettikuviin. Niihin lukeutuvat muun muassa haamukuvioita vaihekoodaussuunnassa aiheuttava potilaan liike, pyöristysvääristymä, aliasoituminen eli liian pienen kuvausalan käytöstä aiheutuva ylitysvääristymä, paikallisen magneettikentän pieni kemiallinen siirtymä, kohinasta johtuva taajuusakselin nollalinjavääristymä, kuvan epäsymmetrinen kirkkaus, gradienttivirran muuttumisesta syntyvät vääristymät sekä metallisista kohteista aiheutuvat vääristymät. Osa vääristymistä johtuu ulkoisista tekijöistä, osa niistä on seurausta magneettikuvantamislaitteiden ominaisuuksista. [71]

Fantomit ja simulaatiot mahdollistavat tekstuurivertailun eri kuvantamislaitteiden välillä. Esimerkiksi kirroottista maksaa esittävä malli voidaan valmistaa polystyreenipalloista ja agarosiliuoksesta. Jirák ym. ovat havainneet tällaisen fantomin tekstuurin säilyvän pitkään vakaana [12].

4.1 Signaali-kohinasuhde

Digitaalisten kuvien sisältämä kohina on seurausta satunnaisesta tai stokastisesta prosessista [72]. Eräs hyvin tärkeä kuvanlaadun mitta on signaalitehon S ja kohinatehon N osamäärä, signaali-kohinasuhde SNR :

$$SNR = \frac{S}{N} \quad (32)$$

Lääketieteellisiin kuviin luontaisesti sisältyvä kohina vaikeuttaa niiden tulkintaa. Magneettikuvat ovat kohinaisia useista eri syistä johtuen. Niihin lukeutuvat muun muassa

lämpöliike, magneettikuvantamislaitteiden epätäydellisyydet ja kuvanmuodostamisvaiheessa ilmenevät tekijät sekä potilaan tahattomat liikkeet. Myös vastaanottimen kaistanleveydellä, kuvausalalla, matriisikoolla, kuvatiedonkeruukertojen määrällä, toistoajalla (time of repetition, TR), kaikuajalla (time to echo, TE), poikkeutuskulmalla, magneettikentän voimakkuudella ja radiotaajuuskelojen (radio frequency, RF) ominaisuuksilla on vaikutusta magneettikuvan signaali-kohinasuhteeseen. [30][73]

Magneettikuvantamislaitteen vastaanottimen kaistanleveydellä on suuri merkitys saavutetun signaali-kohinasuhteen kannalta. Kapeaa kaistanleveyttä käytettäessä signaali-kohinasuhde on suurempi, mutta samalla käytössä olevien leikkeiden lukumäärä jää pienemmäksi ja esimerkiksi potilaan tahaton liikkuminen vaikuttaa kuviin voimakkaammin. [73]

Magneettikuvan signaali-kohinasuhteella on yhteys resoluutioon ja moniin muihin kuvanlaatuun vaikuttaviin tekijöihin [74]. Signaali-kohinasuhde riippuu muun muassa vokselitilavuudesta [75-76]. Signaali-kohinasuhde heikkenee vokselitilavuuden pienentyessä, joten korkean resoluution magneettikuvissa ja ohuissa kuvaleikkeissä signaali-kohinasuhde on tyypillisesti matala. Osittaistilavuusvaikutuksella on suurempi merkitys paksuissa leikkeissä. Leikevälien käyttö parantaa signaali-kohinasuhdetta. Käytännön sovelluksissa leikevälit ovat yleensä noin 25-50 % leikepaksuudesta. [73]

Magneettikuvassa näkyvän kudoksen signaali-kohinasuhdetta arvioidaan perinteisesti vertaamalla kudoksen keskimääräistä signaali-intensiteettiä kuvan taustasta mitattujen intensiteettien keskihajontaan. Koska kohinaisen magneettikuvan intensiteetit noudattavat Rice-jakaumaa, voi tämä suoraviivainen lähestymistapa johtaa kohinatason merkittävään aliarviointiin ja edellyttää korjaustermin käyttöä. [77]

Mayerhoeferin ym. mukaan signaali-kohinasuhteen yhteys tekstuurianalyysin luokittelutarkkuuteen on epälineaarinen [78]. Yhteisesiintymämatriisiparametrit eivät ole yhtä alttiita kohinalle kuin gradienttiperusteiset tekstuuriparametrit. Ne ovat kuitenkin myös vähemmän herkkiä havaitsemaan pieniä mutta todellisia muutoksia kuvan intensiteetissä. [79]

Schad ja Lundervold arvioivat resoluution ja signaali-kohinasuhteen kuuluvan tärkeimpiin magneettikuvien tekstuuriin vaikuttaviin tekijöihin [80]. Schad ja Lundervold ovat myös havainneet resoluution ja tekstuuriparametriarvojen välisen suhteen olevan monimutkainen. Se riippuu muun muassa tutkittavasta kudoksesta ja on altis kohinalle. He kuvasivat aivoja ja useita erilaisia fantomeja FLASH-sekvenssillä (fast low angle shot), käyttivät kolmen tekstuuriparametrin yhdistelmiä ja havaitsivat luokittelutarkkuuden paranevan signaali-kohinasuhteen kohentumisen myötä. Luokitteluun parhaiten soveltuvat

parametrit kuitenkin vaihtuivat signaali-kohinasuhteen muuttuessa. Schad ja Lundervold totesivat kudoksen todellisten tekstuuriominaisuuksien nousevan esiin signaali-kohinasuhteen ylittäessä tietyn kynnyksarvon. Signaali-kohinasuhteen parantumisella ei enää tämän jälkeen ollut suurta vaikutusta tekstuuriparametrien arvoihin.

4.2 Leikepaksuus ja resoluutio

Magneettikuvan leikepaksuus ja resoluutio voivat vaikuttaa tekstuurianalyysin tuloksiin sekä suoraan että esimerkiksi signaali-kohinasuhteen välityksellä. Resoluution vaikutusta tekstuurianalyysin tarkkuuteen on tutkittu selvästi leikepaksuutta enemmän.

Mayerhoefer ym. [78] arvioivat yhteisesiintymä- ja jaksonpituusmatriisiparametrien olevan muita tekstuuriparametreja käyttökelpoisempia silloin, kun kuvan resoluutio on tutkittavan kohteen pienimpien fysikaalisten yksityiskohtien suuruusluokkaa. Heidän tutkimustulostensa perusteella yhteisesiintymämatriisiparametrit kuuluvat vakaimpiin ja joustavimpiin tekstuuriparametreihin.

Schadin ja Lundervoldin tutkimuksessa varsinkin yhteisesiintymämatriisiparametrien arvot muuttuivat resoluution parantuessa ja vokselitilavuuden kasvaessa. Edellä mainituista tekijöistä aiheutunut signaali-kohinasuhteen heikkeneminen saattoi osaltaan lisätä tekstuurianalyysin tulosten hajontaa. Schad ja Lundervold havaitsivat esimerkiksi $0,55 \text{ mm} \times 0,55 \text{ mm}$ resoluution riittävän paljastamaan aivojen harmaan aineen tekstuurin, mutta ei välttämättä valkean aineen tekstuuria. [80]

Guggenbuhl ym. mittasivat yhteisesiintymä- ja jaksonpituusmatriisiparametriarvojen toistettavuutta luista otetuista CT-kuvista (computed tomography) vaihtelevilla leikepaksuuksilla (1 mm, 3 mm, 5 mm ja 8 mm). He havaitsivat useimpien jaksonpituusmatriisiparametrien arvojen pienenevän tilastollisesti merkitsevällä tavalla leikepaksuuden kasvun myötä. Yhteisesiintymämatriisiparametrien arvonmuutosten vaihtelu ei ollut yhtä selväpiirteistä. [81]

Jirák ym. tutkivat resoluution vaikutusta erikokoisista polystyreenipalloista ja agarosigeelistä koostuvien fantomien tekstuurin luokitteluun. Fantomien T1- (1700 ms) ja T2-relaksaatioajat (90 ms) vastasivat tyypillisiä biologisista kudoksista mitattuja arvoja. Polystyreenipallojen halkaisijat vaihtelivat 0,8 millimetristä 3,2 millimetriin. Lineaarilla erotteluanalyysillä fantomit onnistuttiin luokittelemaan virheettömästi parasta ($0,43 \text{ mm} \times 0,45 \text{ mm}$) tai toiseksi parasta ($0,77 \text{ mm} \times 0,90 \text{ mm}$) resoluutiota käytettäessä. Resoluution ollessa $1,53 \text{ mm} \times 1,80 \text{ mm}$ oli virheiden määrä 3-8 %. Tällöin pikselit olivat osapuulleen samaa kokoluokkaa toiseksi pienimpien polystyreenipallojen kanssa. [12]

Mayerhoefer ym. mittasivat fantomeista otettujen 3,0 T magneettikuvien interpoloinnin vaikutusta tekstuurianalyysin tarkkuuteen [82]. Tekstuurianalyysi osoittautui virhealttiiksi menetelmäksi tilanteissa, joissa T2-painotteisten kuvien resoluutio oli hyvin matala. Keskimääräisen resoluutiotarkkuuden magneettikuvia interpoloitaessa koheni varsinkin jaksonpituusmatriisiparametrien, osin myös yhteisesiintymä- ja gradienttiparametrien luokittelutarkkuus. Jos resoluutio oli jo valmiiksi korkea, oli luokittelutarkkuus hyvä, eikä interpolointi tuottanut merkittävää lisähyötyä. Mayerhoefer ym. kehottivatkin harkitsemaan matalaresoluutioisten magneettikuvien interpolointia ennen tekstuurianalyysia.

Brynnolfsson ym. tutkivat ADC-painotteisiin diffuusiotensorikuviin sovelletun tekstuurianalyysin herkkyyttä kohinan ja resoluution suhteen. Heidän tutkimuksessaan herkkyys vaihteli parametrikohteisesti. Tutkimustulostensa perusteella Brynnolfsson ym. suosittelivat resoluutioltaan ja kohinatasoltaan samankaltaisten kuvien käyttöä tekstuurianalyysissa. [83]

4.3 Magneettikuvantamisen sekvenssien ja parametrien valinta

Magneettikuvantamisen sekvenssien ja parametrien valinta vaikuttaa olennaisesti sekä magneettikuvien laatuun että niiden lääketieteelliseen käyttökelpoisuuteen. Käytetyillä kuvantamisparametreilla on vaikutusta myös tekstuurianalyysin parametrivalintaan sekä analyysin tarkkuuteen, luotettavuuteen ja toistettavuuteen.

T1-painotteisissa, lyhyiden TE- ja TR-aikojen magneettikuvissa erilaiset T1-relaksaatioajat omaavien kudosten kontrasti on suuri. Lyhyen T1-ajan kudokset ovat kirkkaita T1-painotteisissa kuvissa, koska niiden pitkittäissuuntainen magnetisaatio ehtii palautua lähes kokonaan ennen seuraavaa virittymistä. T1-painotteinen kuvantaminen on käyttökelpoinen menetelmä varsinkin anatomisten rakenteiden tutkimisessa. Ihmiskudosten T1-palautumisaika on yleensä noin 500-5000 ms. Rasvakudoksen T1-aika on selvästi tätä lyhyempi (noin 260 ms). Myös aivojen valkean aineen T1-aika on suhteellisen lyhyt (noin 780 ms), samoin harmaan aineen T1-aika on lyhyehkö (noin 920 ms). Esimerkiksi veden sekä aivo- ja selkäydinnesteen T1-aika on noin 4000 ms. [73]

T2-painotteisissa, pitkien TE- ja TR-aikojen magneettikuvissa näkyvät kirkkaina kudokset, joiden T2-relaksaatioaika on pitkä. TE-ajan merkittävä pidentäminen johtaa kuitenkin signaali-kohinasuhteen heikkenemiseen. Ihmiskudosten tavanomainen T2-aika on noin 100-300 ms, tosin rasvakudoksen T2-aika on vain noin 80 ms ja valkean aineen T2-aika noin 90 ms. Harmaan aineen T2-aika on noin 100 ms. Veden sekä aivo- ja selkäydinnesteen T2-aika on jopa noin 2000 ms. T2-painotteisia kuvia hyödynnetään usein esimerkiksi aivosairauksia diagnosoitaessa. [73]

Mayerhoeferin ym. mukaan osa tekstuuriparametreista on erityisen herkkiä magneettikuvantamisparametrien arvonmuutoksille [78]. He kuvasivat fantomeja ja mittasivat tekstuuriparametrien arvonvaihtelun suhteessa käytettyihin magneettikuvantamisprotokoliin, TR- ja TE-parametreihin sekä resoluutioon. Tekstuuriparametrit tulivat herkemmiksi magneettikuvantamisparametrien valinnalle resoluution kasvaessa.

Waugh ym. kuvasivat agaroosiliuksesta ja neljän erilaisen karkeustason vaahdosta koostuvia fantomeja usealla eri magneettikuvaussekvenssillä sekä 1,5 T ja 3,0 T kentänvoimakkuuksilla [84]. Tämän jälkeen he arvioivat signaali-kohinasuhteen vaikutusta tekstuurianalyysin tarkkuuteen. Vaikka signaali-kohinasuhde vaihteli sekvensseittäin, ei selkeää yhteyttä mitatun signaali-kohinasuhteen ja luokitteluvirheiden määrän välillä löytynyt. Yhteisesiintymämatriisiparametrien käyttö mahdollisti tarkat luokittelutulokset varsinkin 3,0 T kentänvoimakkuutta ja korkeaa resoluutiota käytettäessä. Waugh ym. kehittivät kiinnittämään erityistä huomiota resoluution merkitykseen tekstuurianalyysia sovellettaessa. Waughin ym. tutkimuksessa varsinkin magneettikuvauksen TR-parametrivalinnalla oli vaikutusta tekstuuriin. TE-parametrivalinnan merkitys oli vähäisempi.

Ford ym. ovat havainneet digitaalisesti simuloitujen magneettikuvien tekstuurin vaihtelevan merkittävästi magneettikuvantamissekvenssistä ja kuvantamisparametreista riippuen [85]. He tutkivat useita erilaisia T1-painotteisia sekvenssejä, mukaan lukien spin-kaikusekvenssi (spin echo, SE), kenttäkaikusekvenssi (gradient echo, GE), käänteispa-lautumisspinkaikusekvenssi (inversion recovery spin echo, IR-SE) ja käänteispa-lautumiskenttäkaikusekvenssi (inversion recovery gradient echo, IR-GE). Fordin ym. tutkimien tekstuuriparametrien arvot vaihtelivat simuloituissa magneettikuvissa muutamasta prosentista jopa yli sataan prosenttiin.

Mayerhoefer ym. sovelsivat tekstuurianalyysia T1- ja T2-painotteisiin, maksakystia ja verisuoniluomia sisältäviin 3,0 T magneettikuviiin [86]. Luokittelutarkkuus oli suurempi T2-painotteisissa kuvissa, vaikka T1-painotteisissa kuvissa oli parempi resoluutio. Kystat ja verisuoniluomet pystyttiin erottelemaan toisistaan 82-88 % tarkkuudella LDA- tai k-NN-algoritmeja käyttämällä.

Herlidou-Même ym. tutkivat tekstuurimenetelmien yleistettävyyttä analysoimalla kolmessa eri tutkimuskeskuksissa otettuja kuvasarjoja [87]. He löysivät jokaisessa tutkimuskeskuksessa otetuista magneettikuvista tekstureja, jotka poikkesivat aivokasvaimista kärsineiden potilaiden ja kontrolliryhmän välillä keskuksissa käytettyjen kuvantamisprotokollien eroavaisuuksista huolimatta. Sitä vastoin Fruehwald-Pallamar ym. suhtautuivat kriittisesti eri tutkimuskeskuksissa otettujen kuvasarjojen tekstuurien hyö-

dyntämiseen, vaikka pitivätkin tekstuurianalyysia käyttökelpoisena kudoslukittelumenetelmänä [88]. Larroza ym. ovat tosin huomauttaneet, ettei Fruehwald-Pallamarin ym. tutkimuksessa ole nähtävästi tehty yleisesti tarpeellisena pidettyä intensiteettinormalisointia ennen tekstuurianalyysia [1]. Mainittu metodologinen puute on saattanut vaikuttaa merkittävästi Fruehwald-Pallamarin ym. johtopäätöksiin.

4.4 Kontrasti

Kuten edellä on todettu, voidaan magneettikuvien kontrastia muuttaa TR- ja TE-aikojen valinnalla. Sitä on mahdollista kohentaa myös varjoaineilla, joiden käyttö on yleistynyt nopeasti. Esimerkiksi pieninä pitoisuuksina käytettynä T1-aikaa lyhentävät paramagneettiset gadoliniumyhdisteet ovat yleisiä, T1-painotteisen magneettikuvan kontrastia parantavia varjoaineita [89]. Niiden käyttö ei kuitenkaan ole täysin riskitöntä ja voi aiheuttaa osalle tutkittavista henkilöistä vakavia sivuvaikutuksia.

Mahmoud-Ghoneim ym. tarkastelivat magneettikuvien dynaamisen skaalan eli käytössä olevien harmaasävyjen lukumäärän vaikutusta yhteisesiintymämatriisiparametreihin perustuvan valkean aineen alueiden tekstuuriluokittelun tarkkuuteen [90]. Heidän mukaansa yhteisesiintymämatriisiparametrien käytön näkökulmasta on olemassa optimaalinen dynaaminen harmaasävyjen lukumäärä N , joka oli kyseisessä tutkimuksessa 128. Jos harmaasävyyskaala oli liian laaja, tuli yhteisesiintymämatriiseista liian harvoja ja luokittelutuloksista sen vuoksi epätarkkoja. Mahmoud-Ghoneim ym. raportoivat erityisesti tekstuurianalyysin sensitiivisyyden heikentyneen mainitussa tapauksessa. Lisäksi tekstuurianalyysi vaati hyvin pitkän laskenta-ajan harmaasävyyskaalan ollessa tarpeettoman suuri. Laskenta-ajan merkitys korostuu varsinkin päivittäisessä potilastyössä.

4.5 Magneettikuvan intensiteettivääristymät

Magneettikuvan intensiteettivääristymät voivat johtaa virheisiin kudosten luokittelussa ja edellyttävät siten korjausta [91]. Intensiteettivääristymien syihin lukeutuvat muun muassa staattisen magneettikentän epähomogeenisuus ja magneettikuvantamislaitteen kelan epätäydellisyydet [92]. Sekä kuvantamislaitteen RF-ominaispiirteet että lähettävät ja vastaanottavat kelat voivat tuottaa keinotekoisia tekstuuria magneettikuviiin. Yksittäisten kelojen ominaisuuksilla on vaikutusta kuvan muodostumiseen. Kuvan keskiosan ja reunojen tekstuurissa voi tämän vuoksi esiintyä hieman vaihtelua. [80]

Materka ym. ovat tutkineet kuvapikselien intensiteettien keskiarvon ja varianssin muutosten vaikutusta tekstuuriparametreihin [93]. Osa tekstuuriparametreista oli Materkan

ym. mukaan erityisen herkkiä intensiteettien keskiarvon tai varianssin muutoksille. Tätä riippuvuutta voidaan pienentää esikäsittelmällä magneettikuvat tekstuurianalyysin yhteydessä yleisesti käytetyllä 3σ -intensiteettinormalisoinnilla.

Schadin mukaan kuvien tavanomainen esikäsittely segmentointia varten sisältää epähomogeenisuuskorjauksen sekä intensiteettien ja kontrastin normalisoinnin [94]. Collewet ym. tutkivat juustonäytteistä otettujen magneettikuvien kuvantamisprotokollien ja harmaasävynormalisointimenetelmien vaikutusta tekstuurianalyysin tarkkuuteen [95]. Ellei normalisointia ollut tehty, riippui luokitteluvirhe käytetystä protokollasta. Luokittelutarkkuus parani selvästi, kun kuvien intensiteetit korjattiin vierekkäisten pikselien harmaasävyeroja korostavalla 3σ -intensiteettinormalisoinnilla.

Toisaalta Materka ja Strzelecki ovat varoittaneet tutkijoita siitä, ettei 3σ -intensiteettinormalisointi ota huomioon intensiteetin paikallisia vaihteluita [91]. Heidän mukaansa kuvan intensiteettien epätasaisuus voi toisinaan vaikuttaa varsinaista tekstuuria enemmän kudosten luokitteluun ja johtaa vakuuttavalta näyttäviin mutta virheellisiin tuloksiin. Ratkaisuksi ongelmaan Materka ja Strzelecki ovat ehdottaneet paikallisiin intensiteettikeskiarvoihin perustuvaa polynomista korjausta tavanomaisen 3σ -intensiteettinormalisoinnin lisäksi.

4.6 Tarkasteltavan ROI-alueen valinta

Myös tarkasteltavan ROI-alueen koko ja muoto saattavat vaikuttaa tekstuurianalyysin tarkkuuteen. Harrison ym. eivät ole silti pitäneet ROI-alueen asettelua esimerkiksi MS-taudin tekstuurianalyysin tarkkuuden kannalta ratkaisevana tekijänä [37]. He tutkivat tekstuurianalyysin erottelutarkkuutta MS-plakkien sekä normaalin valkean aineen (normal white matter, NWM) ja normaalilta näyttävän valkean aineen (normal appearing white matter, NAWM) välillä. Neuroverkkoja ja lähin naapuri -luokittelumenelmiä käytettäessä luokittelutarkkuus oli hyvin korkea (vähintään 96 %). Harrisonin ym. käyttämä kuvantamissekvenssi ei vaikuttanut olennaisesti luokittelun tarkkuuteen.

Vaikka vapaasti piirretyt ROI-alueet mahdollistivat hieman tarkemmat luokittelutulokset vakiokokoisiin ROI-valintoihin verrattuna, pitivät Harrison ym. vapaasti piirrettyjen ROI-alueiden käyttöä liian paljon aikaa vievänä ja tuloksia vaikeammin toistettavana [37]. Heidän mukaansa valinta vakiokokoisien ja vapaasti piirrettyjen ROI-alueiden välillä riippuu tutkimuskohteesta ja tilanteen asettamista vaatimuksista. Tutkijat painottivat myös tilanteeseen soveltuvan kuvantamissekvenssin valinnan tärkeyttä.

Sikiö ym. tutkivat ROI-alueiden koon vaikutusta magneettikuvien tekstuurianalyysin tarkkuuteen 64 terveen aikuisen keinotekoista kohinaa sisältäneillä kuvilla [96]. He havaitsivat sen vaikuttavan merkittävästi toisen kertaluvun tekstuuriparametrien arvoihin. Tutkijat päättelivät, että erikokoisten ROI-alueiden käyttö tekstuurianalyysissa voi pahimmillaan johtaa virheellisiin luokittelutuloksiin.

5. AINEISTOT JA MENETELMÄT

Tämän työn tarkoituksena on tarkastella magneettikuvien leikepaksuuden ja signaali-kohinasuhteen vaikutusta lääketieteellisen tekstuurianalyysin tarkkuuteen ja luotettavuuteen. Työssä myös tutkitaan, voidaanko tekstuurianalyysin tarkkuutta parantaa hyödyntämällä sitä yhdessä diffuusiotensorikuvauksen kanssa. Tutkimusaineisto koostuu MS-potilaiden aivoista ja terveiden vapaaehtoisten tutkimuskohteiden reiden alueesta otetuista magneettikuvista.

Työssä on hyödynnetty useita erilaisia tekstuuriparametreja, joista suurin osa perustuu yhteisesiintymämatriiseihin. Kaikkien tekstuuriparametrien arvot on laskettu vapaasti saatavissa olevalla MaZda-ohjelmalla [97] magneettikuvien 3σ -intensiteettinormalisoinnin jälkeen. DTI-parametriarvojen laskemiseen on käytetty Siemens Trio MR Neuro3D Engine -laskentaohjelmistoa.

Varsinaisessa luokittelussa käytetyt tekstuuriparametrit on valittu Fisherin menetelmällä. Kunkin tekstuuriparametrin F-lukuarvo on laskettu jakamalla tarkasteltujen luokkien välinen varianssi luokkien keskimääräisellä sisäisellä varianssilla. Tämän jälkeen luokitteluun on valittu ne parametrit, joiden F-lukuarvot ovat olleet korkeimmat. Fisherin menetelmää pidetään yleisesti luotettavana parametrivalintamenetelmä ja niin sanottuna ”kultaisena standardina” monissa tekstuurianalyysisovelluksissa [37][98-99].

Tutkittavat kudokset on parametrien laskemisen ja valinnan jälkeen luokiteltu useilla erilaisilla algoritmeilla ristiinvalidointia apuna käyttäen. Yksittäisissä luokitteluissa on käytetty erillisiä opetus- ja testiaineistoja, jotta tuloksista ja niiden yleistettävyydestä saataisiin mahdollisimman todenmukainen käsitys.

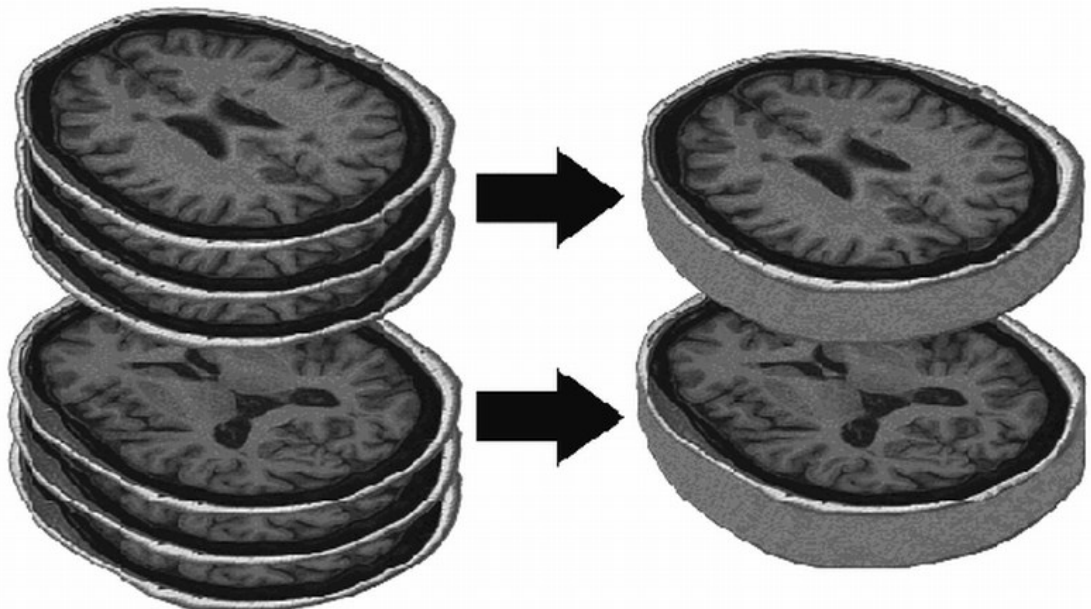
Luvuissa 5.1, 5.2 ja 5.3 kuvatut tutkimukset on julkaistu aiemmin itsenäisinä tutkimusartikkeleina ([I], [II] ja [III], Osatyöt I-III). Tutkimustyön kannalta olennaisimmat magneettikuvausparametrit on esitetty Taulukossa 1.

Taulukko 1. Yhteenveto Osatyöissä I-III ([I-III]) käytetyistä magneettikuvausparametreista. Osatyön III diffuusiotensorikuvauskohtaiset parametrit on esitetty sulkeissa.

Kuvantamisparametri	Osatyö I (Leikepaksuus)	Osatyö II (SNR)	Osatyö III (TA & DTI)
Kentänvoimakkuus	1,5 T	1,5 T	1,5 T
Leikepaksuus	0,9 mm	3,0 mm	0,9 mm (5,0 mm)
Resoluutio	0,45 mm * 0,45 mm	0,81 mm * 0,81 mm	0,45 mm * 0,45 mm (1,8 mm * 1,8 mm)
Kuvantamissekvenssi	3D MPR GE T1/T1C	Axial T2* GE	3D MPR GE T1
Toisto aika (TR)	1160 ms	36 ms	1160 ms (3500 ms)
Kaikuaika (TE)	4,24 ms	15 ms	4,24 ms (96 ms)
Inversio aika (TI)	600 ms	N/A	600 ms
Poikkeutuskulma (°)	15	12	15

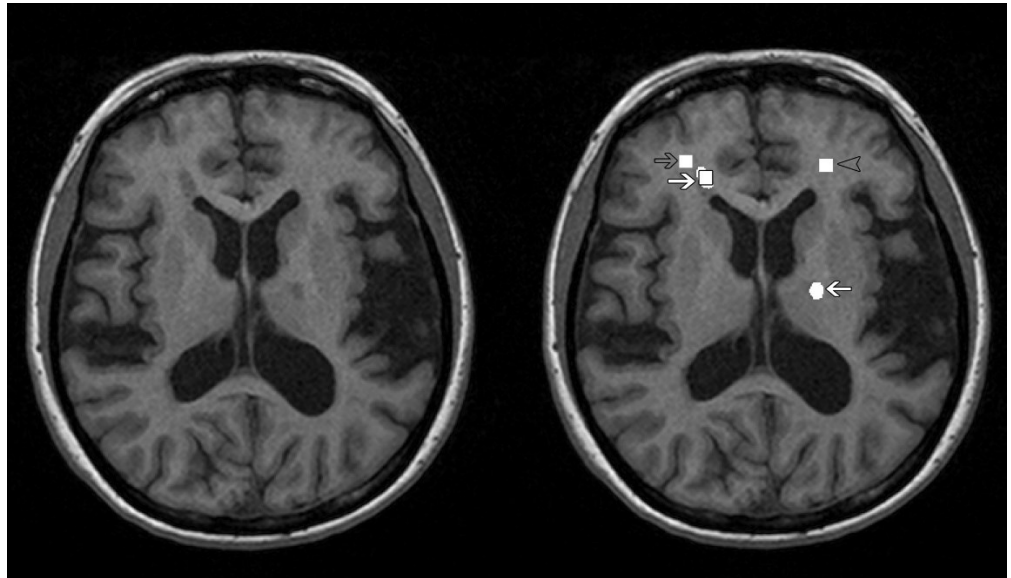
5.1 Leikepaksuuden vaikutus tekstuurianalyysin tarkkuuteen

Leikepaksuuden vaikutusta tekstuurianalyysin tarkkuuteen on tutkittu Osatyössä I [I]. Tutkimusaineistona on käytetty MS-tautia sairastavien potilaiden aivoista T1-painotteisella sekvenssillä otettuja magneettikuvia. Lisäksi osassa T1-painotteisia kuvantamistutkimuksia hyödynnettiin kontrastia parantavaa, gadolinium-pohjaista varjoainetta (T1 with contrast agent, T1C). Kuvien alkuperäinen leikepaksuus oli noin 1 mm, ja simuloidut 3 mm leikkeet muodostettiin kolmen peräkkäisen 1 mm leikkeen intensiteettikeskiarvoina kahdella aivoalueella (Kuva 3).



Kuva 3. Kolmen peräkkäisen 1 mm magneettikuvaleikkeen (vasemmalla) pikseli-intensiteetit keskiarvoistettiin simuloiduksi, 3 mm paksuksi leikkeeksi (oikealla). Keskiarvoistuksia suoritettiin kahdella aivoalueella. [I]

Alkuperäisistä ja simuloituista leikkeistä laskettiin 3σ -intensiteettinormalisoinnin jälkeen 264 tekstuuriparametrin arvot valkeassa aineessa, normaalilta näyttävässä valkeassa aineessa ja MS-plakkien alueella. Tutkitut plakit merkittiin kuviin sekä vakiokokoisilla, 10×10 pikselin (20 mm^2) suuruisilla (MS regular, MSr) että vapaasti piirretyillä, $39\text{--}2477$ pikselin ($8\text{--}502 \text{ mm}^2$) kokoisilla (MS irregular, MSi) ROI-alueilla (Kuva 4).

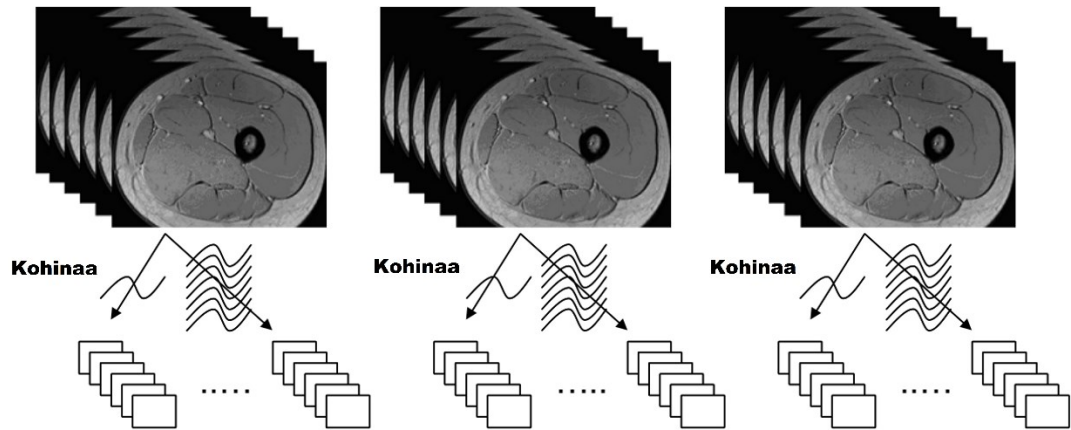


Kuva 4. Tutkitut ROI-alueet. Vasemmalla: Alkuperäinen T1-painotteinen magneettikuva MS-potilaasta. Oikealla: Esimerkki ROI-alueista. Normaalilta näyttävä valkea aine on merkitty kuvaan harmaalla nuolella ja MS-plakit (vakiokokoinen tai vapaasti piirretty ROI) valkoisilla nuolilla. Normaali valkea aine on merkitty harmaalla nuolenpäällä. [1]

Valkean aineen ja MS-plakkien välisiä tekstuurieroja analysoitiin Wilcoxonin merkittyjen sijalukujen testillä. LDA- ja NDA-luokitteluissa käytetyt tekstuuriparametrit valittiin tämän jälkeen Fisherin menetelmällä. Valkean aineen alueiden ja MS-plakkien luokittelua varten muodostettiin erilliset opetus- ja testiaineistot. Ristiinvalidoinnin avulla toteutetussa opetuksessa ja testauksessa käytettiin sekä alkuperäisiä että simuloituja leikkeitä. Opetus- ja testausprosessi toistettiin jokaisen tarkastellun kudosparin osalta 100 kertaa, minkä jälkeen laskettiin keskimääräinen luokittelutarkkuus ja tulosten hajonta. [1]

5.2 Signaali-kohinasuhteen vaikutus tekstuurianalyysin tarkkuuteen

Signaali-kohinasuhteen vaikutusta tekstuurianalyysin tarkkuuteen tutkittiin Osatyössä II [II]. Vapaaehtoisten terveiden tutkimuskohteiden reisilihaksista otettuihin 18 alkuperäiseen magneettikuvaan lisättiin tutkimusta varten keinotekoisia, normaalijakautunutta kohinaa magneettikuvainformaation taajuus- ja vaihemuodossa sisältävässä mitausvaruudessa eli k-avaruudessa (Kuva 5).



Kuva 5. Keinotekoisista kohinaa sisältävien 126 kuvan simulointi 18 alkuperäisestä magneettikuvasta. Kustakin alkuperäisestä kuvasta luotiin kohinan määrää asteittain lisäämällä seitsemän simuloitua kuvaa. Sekä alkuperäisiä että simuloituja kuvia hyödynnettiin tekstuurianalyysissa. [II]

Keinotekoisien kohinan lisäämisen jälkeen käytössä oli yhteensä 144 alkuperäistä tai simuloitua magneettikuvaa, joista pyrittiin erottelemaan tekstuurianalyysillä neljä eri kudosta (suora reisilihas (m. rectus femoris) ja ulompi reisilihas (m. vastus lateralis), luuydin sekä rasva). Rasvaan ja lihaskudokseen merkittiin neliönmuotoiset, 144 pikselin (94 mm^2) suuruiset ROI-alueet. Luuytimeen piirretyt ympyränmuotoiset ROI-alueet olivat 68 pikselin (45 mm^2) kokoisia. (Kuva 6)



Kuva 6. (A) Alkuperäinen reiden alueen magneettikuva. (B) Simuloitua kohinaa sisältävä magneettikuva, jossa näkyvät reiden kahteen lihakseen (suora reisilihas (I) ja ulompi reisilihas (II)) sekä luuytimeen (III) ja rasvakudokseen (IV) merkityt ROI-alueet. [II]

Kuvat jaettiin signaali-kohinasuhteen perusteella kuuteen 24 kuvan ryhmään. Niiden signaali-kohinasuhteet vaihtelivat kudostyyppistä riippuen välillä 0-15 dB (SNR-luokka 1), 7-20 dB (SNR-luokka 2), 15-27 dB (SNR-luokka 3), 21-32 dB (SNR-luokka 4), 30-39 dB (SNR-luokka 5) ja 37-47 dB (SNR-luokka 6) (Taulukko 2).

Taulukko 2. Signaali-kohinasuhteen (dB) ala- ja ylärajat kudiskohtaisille (suora reisilihas, ulompi reisilihas, rasva ja luuydin) SNR-luokille 1-6. [II]

SNR-luokka	Suora reisilihas	Ulompi reisilihas	Rasva	Luuydin
1	1-13	3-15	0-11	0-6
2	14-18	16-20	12-18	7-14
3	19-24	21-27	19-24	15-20
4	25-30	28-32	25-32	21-29
5	31-36	33-37	33-39	30-36
6	37-41	38-42	40-47	37-43

Kunkin kuvan neljästä ROI-valinnasta laskettiin 3σ -intensiteettinormalisoinnin jälkeen 220 yhteisesiintymämatrisiparametria, 20 jaksonpituusmatrisiparametria, 5 autoregressiivista parametria, 4 gradienttiparametria ja 4 wavelet-parametria. Lopuksi kudokset luokiteltiin kolmen suurimman F-lukuarvon saaneen tekstuuriparametrin avulla LDA-, 1-NN- ja 3-NN-menetelmiä hyödyntäen. Opetus ja varsinainen luokittelu toteutettiin tässäkin tutkimuksessa kunkin kudosparin osalta 100 kertaa keskimääräisen luokittelutarkkuuden ja tulosten hajonnan määrittämiseksi. Luokitteluissa käytettiin opetusaineistona 12 ROI-aluetta SNR-luokasta 5 tai 6 sekä testiaineistona 12 ROI-aluetta SNR-luokista 1-6. [II]

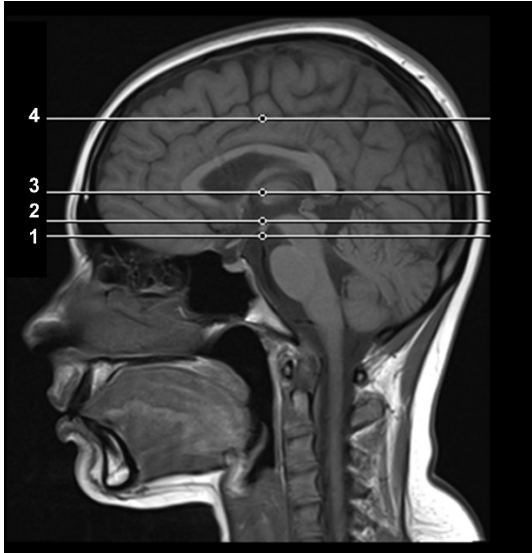
5.3 Tekstuurianalyysin ja diffuusiotensorikuvauksen yhdistäminen

Tekstuurianalyysia ja siihen yhdistettyä diffuusiotensorikuvausta tutkittiin Osatyössä III [III] hyödyntämällä 17 PPMS- ja 19 RRMS-potilaan T1-painotteisia magneettikuvia ja niitä vastaavia diffuusiotensorikuvia. Analyysiin valittiin symmetriset alueet aivojen vasemmasta ja oikeasta puoliskosta.

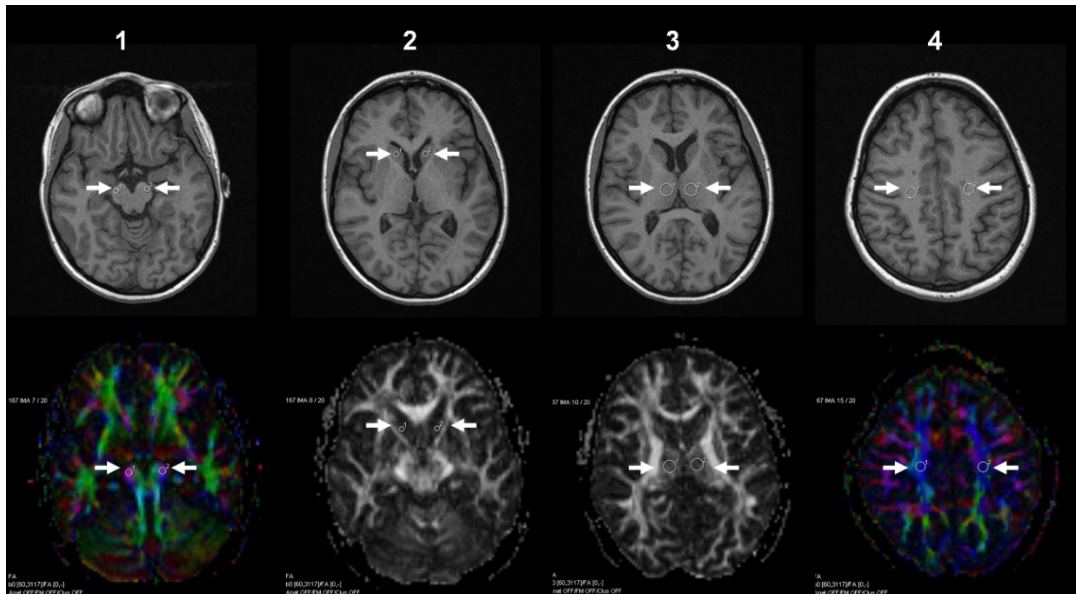
Harmaan aineen alueista tutkimuksen kohteena olivat caudate nucleus ja thalamus sekä valkean aineen alueista cerebral peduncle ja centrum semiovale. Niihin piirrettyjen ympyränmuotoisten ROI-alueiden koko oli 18-82 mm². Lisäksi sekä T1-painotteisiin magneettikuviin että diffuusiotensorikuviin merkittiin 18 mm²:n kokoiset ympyränmuotoiset ROI-alueet MS-plakkien kohdalle ja symmetrisesti toiselle aivopuoliskolle. ROI-valinnan periaatteet on esitetty Kuvissa 7-9.

Magneettikuviin piirretyistä ROI-alueista laskettiin 3σ -intensiteettinormalisoinnin jälkeen 132 tekstuurianalyysiparametrin arvot. Ne koostuivat 11 erilaisesta yhteisesiintymämatrisiparametrin arvosta 0, 45, 90 ja 135 asteen suunnissa sekä 1, 2 ja 3 pikselin etäisyyksillä. DTI-kuviin piirretyistä ROI-alueista mitattiin FA- ja ADC-arvot Siemens Trio MR Neuro3D Engine -laitteella. Lisäksi mitattiin tekstuuriparametrien ja DTI-arvojen parametrikohdalliset korrelaatiot.

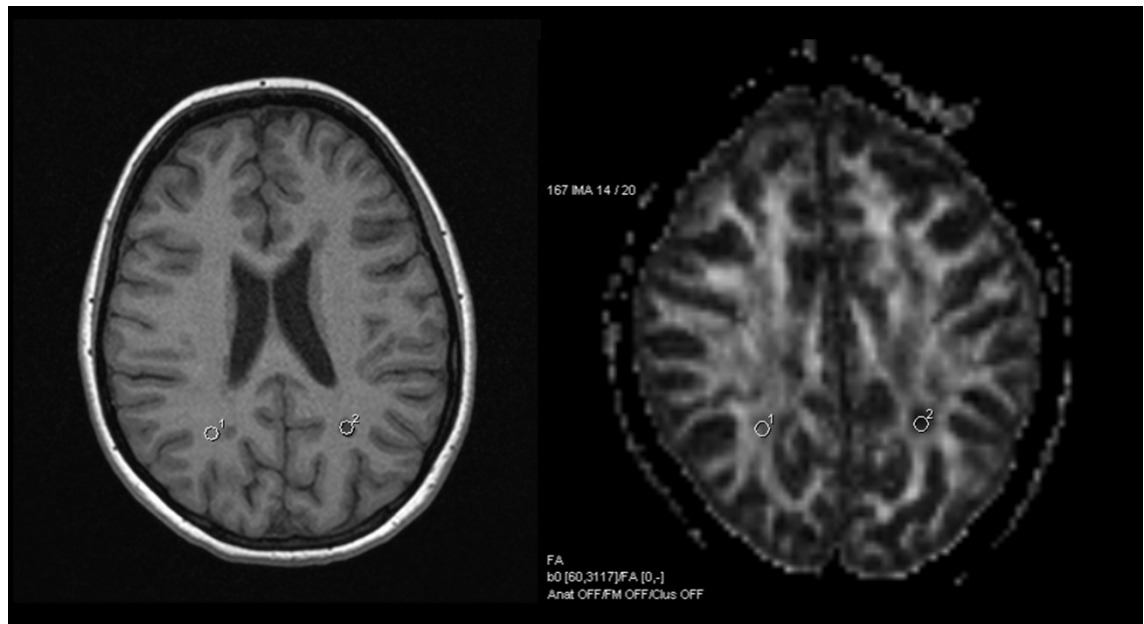
Tekstuuri- ja DTI-parametrien arvoja käytettiin aivojen vasemman ja oikean puoliskon välisten erojen määrittelyyn Wilcoxonin merkittyjen sijalukujen testin avulla. Lopuksi hyödynnettiin kahta korkeimman F-lukuarvon saanutta tekstuuriparametria sekä FA- ja ADC-parametreja PPMS- ja RRMS-tautityyppien luokitteluun LDA-, NDA-, 1-NN- ja 3-NN-menetelmillä. Tuloksia arvioitiin ristiinvalidoinnin keinoin. Menetelmien tehokkuuden mittaamiseen käytettiin tuhatta erilaista opetus- ja testiaineistojen yhdistelmää. [III]



Kuva 7. Esimerkki leikevalinnasta. Tutkitut leiketason: (1) cerebral peduncle, (2) caudate nucleus, (3) thalamus ja (4) centrum semiovale. [III]



Kuva 8. T1-painotteiset magneettikuvat (ylärivi) ja FA-kartat (alarivi). Valkoisilla nuolilla merkityt, ympyränmuotoiset ROI-alueet piirrettiin symmetrisesti molemmille aivopuoliskoille seuraaviin alueisiin: (1) cerebral peduncle, (2) caudate nucleus, (3) thalamus ja (4) centrum semiovale.



Kuva 9. Esimerkki MS-plakin merkitsemisestä RRMS-potilaan T1-painotteisessa magneettikuvassa (vasemmalla) sekä vastaavan ROI-alueen piirtämisestä FA-kartassa (oikealla). Vastaavat ympyränmuotoiset ROI-alueet merkittiin symmetrisesti myös toiselle aivopuoliskolle. [III]

6. TUTKIMUSTULOKSET

6.1 Leikepaksuus

Aivojen magneettikuvien tekstuuria vertailtaessa havaittiin Osatyössä I [1] ainoastaan pienehköjä eroavaisuuksia tutkimuksessa käytettyjen leikepaksuuksien (alkuperäinen 1 mm ja simuloitu 3 mm) välillä (Taulukot 3 ja 4). LDA- ja NDA-menetelmillä toteutettu luokittelu osoittautui 3 mm leikkeillä lähes yhtä tarkaksi kuin 1 mm leikkeillä. Useimmissa tapauksissa luokittelutarkkuus oli vähintään 90 %. Selkeä poikkeus oli valkean aineen ja normaalilta näyttävän valkean aineen luokittelu, jonka tarkkuus vaihteli välillä 48-62 %. Valkean aineen alueet pystyttiin erottelamaan plakeista hyvällä tarkkuudella siinäkin tapauksessa, että opetus- ja testiaineistojen leikepaksuudet olivat erilaisia.

Taulukko 3. LDA-menetelmän tarkkuus (%) T1-painotteisissa (T1) sekä varjoainetta hyödyn-
täen muodostetuissa (T1C) aivokuvissa. Vertailtavina olivat aivojen valkea aine
(WM), normaalilta näyttävä valkea aine (NAWM), vapaasti piirretyllä ROI-alueella
merkityt MS-plakit (MSi) sekä vakiokokoisella ROI-alueella merkityt MS-
plakit (MSr). Jokaisessa vertailussa on käytetty kolmea Fisherin menetelmällä
valittua tekstuuriparametria. Ylärivin merkintä "X -> Y" viittaa tilanteeseen, jossa
opetusaineistoon on valittu X mm (1 mm/3 mm) paksuja leikkeitä ja testiaineis-
toon Y mm (1 mm/3 mm) paksuja leikkeitä. 95 prosentin varmuusväli on esitetty
hakasuluissa kursiivilla. [1]

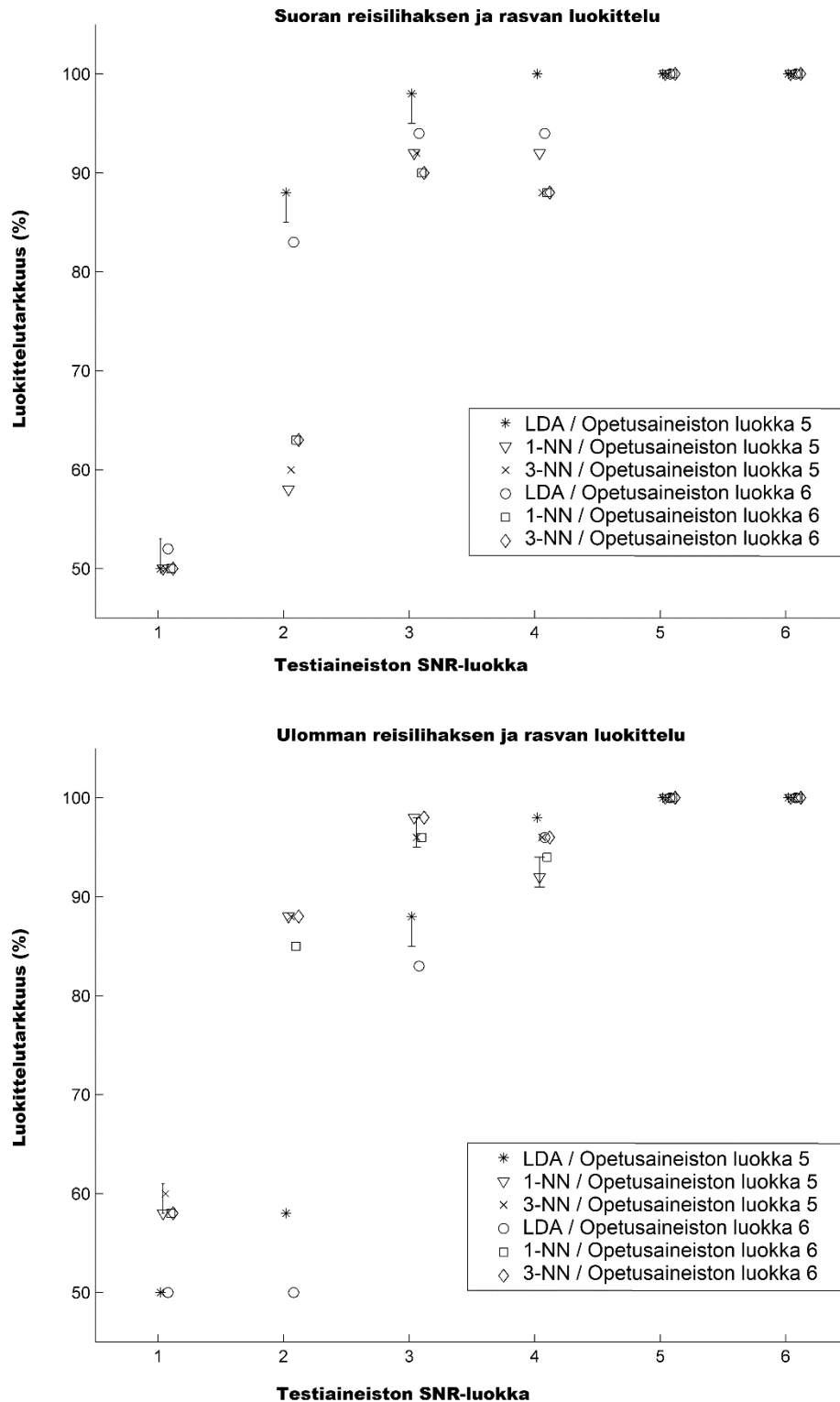
ROI-vertailu	1 mm -> 1 mm		1 mm -> 3 mm		3 mm -> 1 mm		3 mm -> 3 mm	
	T1	T1C	T1	T1C	T1	T1C	T1	T1C
WM vs. NAWM	60 [59, 61]	55 [53, 56]	52 [50, 54]	50 [48, 53]	57 [55, 58]	55 [53, 56]	60 [59, 61]	62 [60, 63]
WM vs. MSi	92 [91, 95]	98 [97, 98]	93 [93, 94]	93 [92, 93]	92 [91, 95]	89 [89, 90]	96 [95, 96]	95 [94, 95]
WM vs. MSr	93 [93, 94]	93 [92, 94]	88 [88, 89]	91 [91, 93]	90 [90, 91]	84 [83, 85]	89 [88, 89]	92 [92, 93]
NAWM vs. MSi	91 [90, 92]	98 [97, 98]	92 [92, 93]	98 [98, 98]	95 [94, 96]	98 [98, 100]	95 [94, 96]	100 [98, 100]
NAWM vs. MSr	91 [90, 92]	90 [89, 91]	88 [86, 88]	90 [89, 90]	90 [90, 91]	91 [90, 92]	88 [86, 88]	94 [93, 94]

Taulukko 4. *NDA-menetelmän tarkkuus (%) T1-painotteisissa (T1) sekä varjoainetta hyödyntäen muodostetuissa (T1C) aivokuvissa. Vertailtavina olivat aivojen valkea aine (WM), normaalilta näyttävä valkea aine (NAWM), vapaasti piirretyllä ROI-alueella merkityt MS-plakit (MSi) sekä vakiokokoisella ROI-alueella merkityt MS-plakit (MSr). Jokaisessa vertailussa on käytetty kolmea Fisherin menetelmällä valittua tekstuuriparametria. Ylärivin merkintä "X -> Y" viittaa tilanteeseen, jossa opetusaineistoon on valittu X mm (1 mm/3 mm) paksuja leikkeitä ja testiaineistoon Y mm (1 mm/3 mm) paksuja leikkeitä. 95 prosentin varmuusväli on esitetty hakasuluissa kursiivilla. [I]*

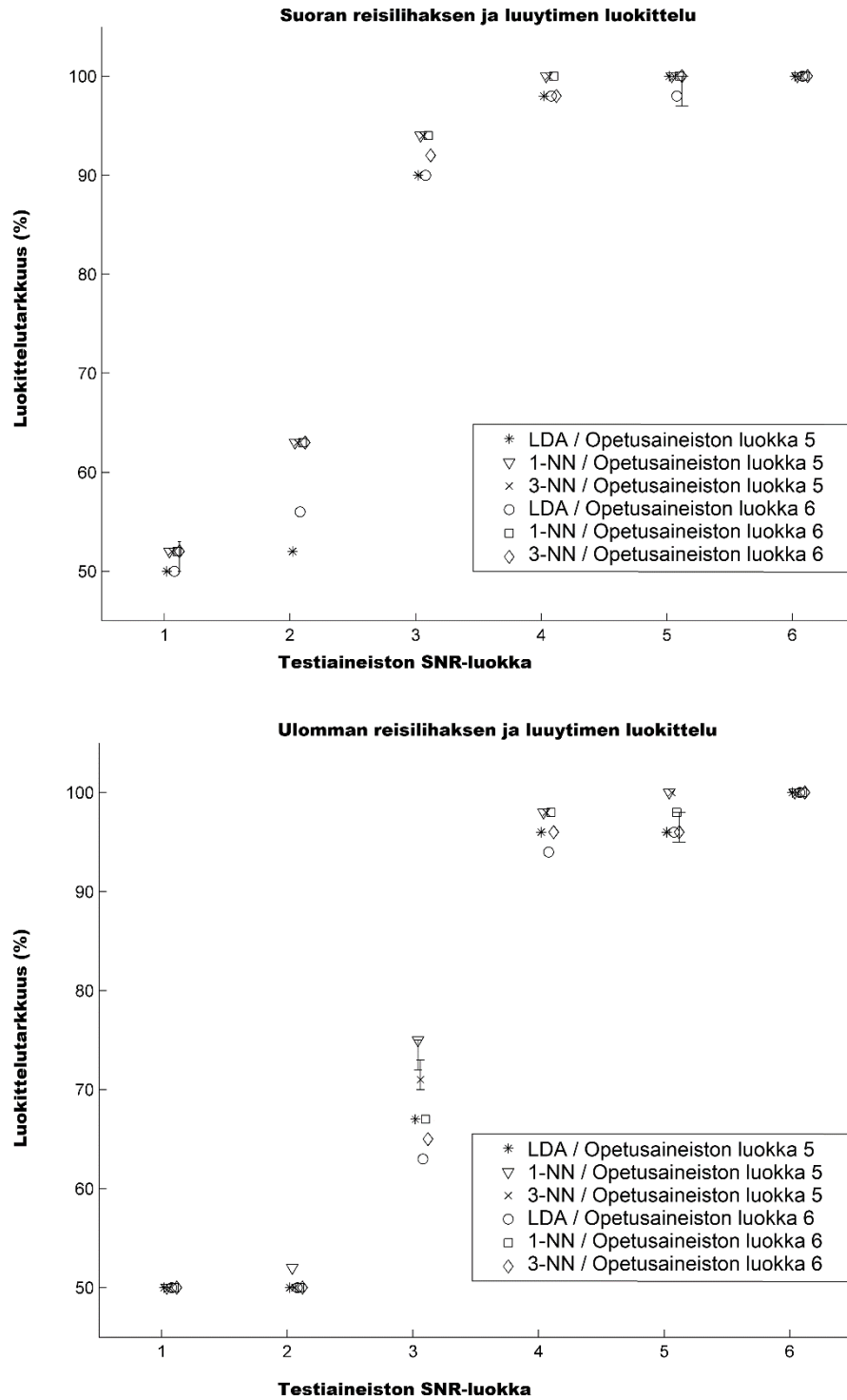
ROI-vertailu	1 mm -> 1 mm		1 mm -> 3 mm		3 mm -> 1 mm		3 mm -> 3 mm	
	T1	T1C	T1	T1C	T1	T1C	T1	T1C
WM vs. NAWM	59 [57, 60]	50 [48, 52]	54 [52, 56]	48 [46, 50]	54 [54, 55]	50 [50, 53]	57 [55, 58]	60 [59, 63]
WM vs. MSi	95 [94, 96]	96 [94, 97]	94 [93, 95]	92 [91, 93]	92 [91, 93]	88 [88, 89]	96 [96, 98]	93 [92, 94]
WM vs. MSr	92 [92, 93]	92 [91, 93]	85 [85, 86]	92 [91, 92]	90 [89, 90]	81 [80, 82]	88 [88, 89]	92 [91, 92]
NAWM vs. MSi	92 [91, 92]	98 [96, 98]	91 [90, 92]	97 [96, 98]	91 [90, 93]	98 [97, 98]	95 [93, 96]	98 [98, 100]
NAWM vs. MSr	90 [89, 90]	88 [88, 89]	86 [85, 88]	90 [89, 91]	89 [88, 90]	90 [89, 91]	86 [85, 86]	92 [91, 92]

6.2 Signaali-kohinasuhde

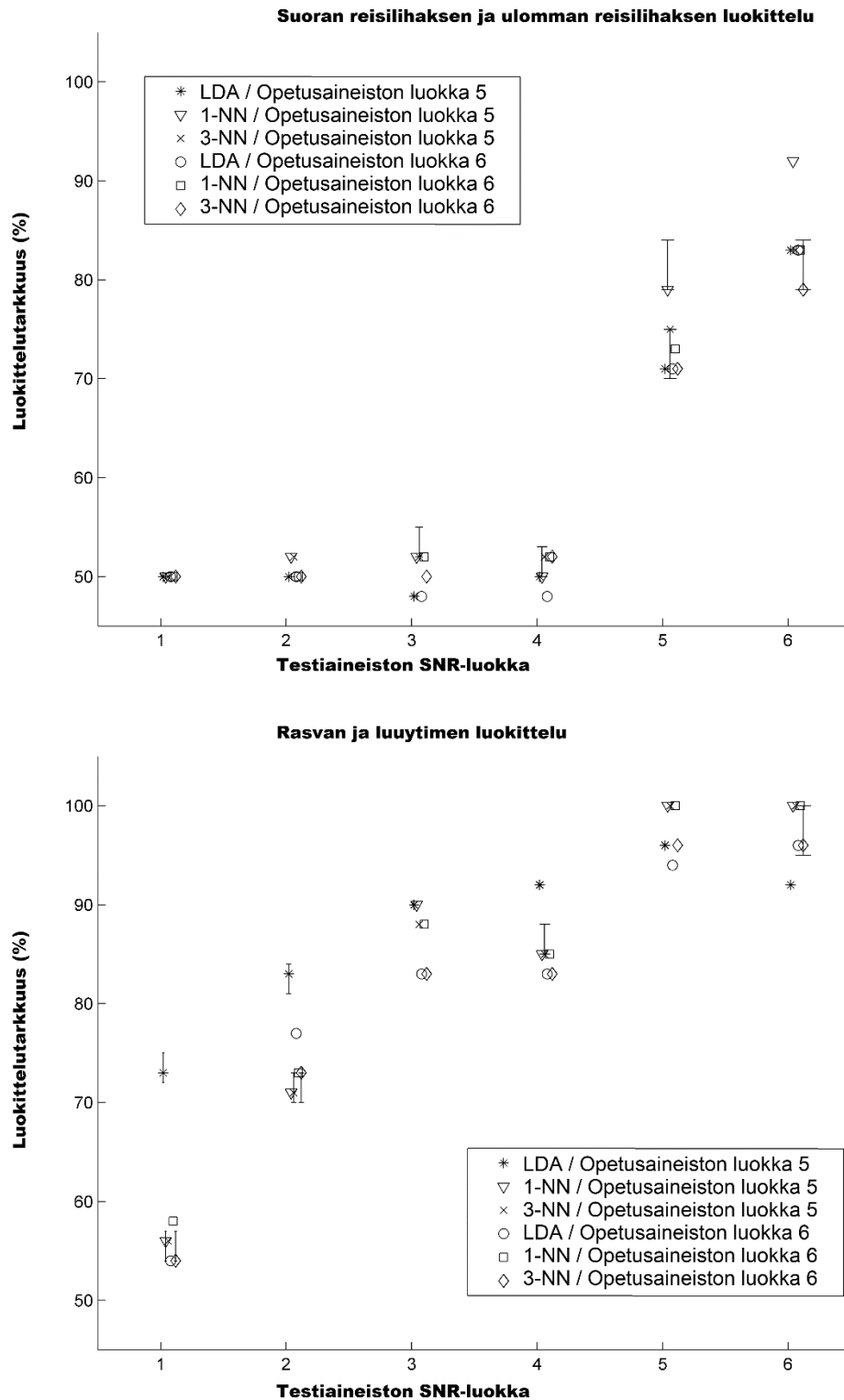
Osatyössä II [II] tutkittiin signaali-kohinasuhteen vaikutusta tekstuurianalyysin erottelutarkkuuteen ja sen hajontaan reiden alueen kudoksia (suora reisilihas, ulompi reisilihas, luuydin ja rasva) luokiteltaessa. Useimmat kudostyyppit erottuivat toisistaan vähintään 90 % tarkkuudella, jos signaali-kohinasuhde oli riittävän korkea. Luokittelutulokset on esitetty Kuvissa 10-12.



Kuva 10. Suoran reisilihaksen (*m. rectus femoris*) ja rasvan (yläkuva) sekä ulomman reisilihaksen (*m. vastus lateralis*) ja rasvan (alakuva) LDA-, 1-NN- ja 3-NN-luokittelut. Vaaka-akselilla on esitetty testiaineiston SNR-luokka sekä pystyakselilla luokittelutarkkuus ja sen mahdollinen (vähintään 3 %-yksikön suuruinen) hajonta. Opetusaineistona on käytetty SNR-luokkia 5-6 ja testiaineistona SNR-luokkia 1-6. [11]



Kuva 11. Suoran reisilihaksen (*m. rectus femoris*) ja luuytimen (yläkuva) sekä ulomman reisilihaksen (*m. vastus lateralis*) ja luuytimen (alakuva) LDA-, 1-NN- ja 3-NN-luokittelut. Vaaka-akselilla on esitetty testiaineiston SNR-luokka sekä pysty-akselilla luokittelutarkkuus ja sen mahdollinen (vähintään 3 %-yksikön suuruinen) hajonta. Opetusaineistona on käytetty SNR-luokkia 5-6 ja testiaineistona SNR-luokkia 1-6. [II]



Kuva 12. Suoran reisilihaksen (*m. rectus femoris*) ja ulomman reisilihaksen (*m. vastus lateralis*) (yläkuva) sekä rasvan ja luuytimen (alakuva) LDA-, 1-NN- ja 3-NN-luokittelut. Vaaka-akselilla on esitetty testiaineiston SNR-luokka sekä pystyakselilla luokittelutarkkuus ja sen mahdollinen 3 %-yksikön suuruinen hajonta. Opetusaineistona on käytetty SNR-luokkia 5-6 ja testiaineistona SNR-luokkia 1-6. [II]

6.3 Tekstuurianalyysin ja diffuusiotesorikuvauksen yhdistäminen

Taulukoissa 5-8 on esitetty yhteenveto Osatyöhön III [III] sisältyvistä, PPMS- ja RRMS-potilaiden vasemman ja oikean aivopuoliskon eroihin liittyvistä tekstuuri- ja diffuusiotesorimittaustuloksista korrelaatioineen sekä MS-tautityyppien luokittelun tulokset. Useimpien tutkimuksessa käytettyjen algoritmien luokittelutarkkuus oli noin 80 %.

Taulukko 5. Vasemman ja oikean aivopuoliskon välillä tilastollisesti merkitsevästi eroavien ($p < 0,01$) tekstuuriparametrien osuus. 132 tekstuuriparametrin p -arvot määritettiin Wilcoxonin merkittyjen sijalukujen testillä 17 PPMS-tautia ja 19 RRMS-tautia sairastavan potilaan neljällä eri aivoalueella (cerebral peduncle, thalamus, caudate nucleus ja centrum semiovale). [III]

Aivojen alue	Tilastollisesti merkitsevästi erilaisten ($p < 0,01$) tekstuuriparametrien osuus (%)	
	PPMS	RRMS
Cerebral peduncle	1	1
Thalamus	1	23
Caudate nucleus	2	1
Centrum semiovale	0	2

Taulukko 6. Wilcoxonin merkittyjen sijalukujen testin osoittamat vasemman ja oikean aivopuoliskon erot diffuusiotesorikuvauksessa. [III]

Aivojen alue	PPMS (n = 17)		RRMS (n = 19)	
	FA	ADC	FA	ADC ($10^{-3} \text{ mm}^2/\text{s}$)
Cerebral peduncle	0,44	0,72	0,38	0,10
Thalamus	0,75	0,30	0,013*	0,39
Caudate nucleus	0,92	0,49	0,006**	0,39
Centrum semiovale	0,94	0,14	0,87	0,025*

** : tilastollisesti merkitsevä ($p < 0,01$)

* : tilastollisesti melkein merkitsevä ($p < 0,05$)

Taulukko 7. GLCM-tekstuuriparametrien ja diffuusiotesoridatan (FA, ADC) välinen mediaanikorrelaatio PPMS- ja RRMS-potilasryhmillä. [III]

GLCM-parametrityyppi	PPMS (n = 17)		RRMS (n = 19)	
	FA	ADC	FA	ADC
Sum of Variance	0,026	0,14	0,14	0,32
Inverse Difference Moment	0,30	0,19	0,055	0,049
Correlation	0,12	0,028	0,15	0,20
Sum of Entropy	0,29	0,27	0,27	0,24
Difference Entropy	0,049	0,023	0,089	0,044
Contrast	0,21	0,049	0,10	0,012
Difference Variance	0,15	0,13	0,14	0,14
Sum of Squares	0,25	0,28	0,15	0,42
Sum Average	0,010	0,29	0,36	0,38
Entropy	0,37	0,35	0,19	0,19
Angular Second Moment	0,45	0,32	0,18	0,16

Taulukko 8. *PPMS- ja RRMS-potilaiden lineaarinen (LDA) ja epälineaarinen (NDA, 1-NN, 3-NN) luokittelu tekstuurianalyysin ja diffuusiotensoridatan avulla. Luokittelussa on käytetty kahta Fisherin menetelmän avulla valittua tekstuuriparametria (keskimäinen sarake) sekä kyseisten tekstuuriparametrien ja FA- ja ADC-arvojen yhdistelmää (oikeanpuoleinen sarake). [III]*

Algoritmi	Luokittelutarkkuus (%) (2 tekstuuriparametria)	Luokittelutarkkuus (%) (2 tekstuuriparametria + FA + ADC)
LDA	73	77
NDA	81	79
1-NN	80	81
3-NN	81	79

7. POHDINTAA

Tämän työn tavoitteena on ollut tutkia lääketieteellisiin magneettikuviin sovelletun tekstuurianalyysin tarkkuutta ja luotettavuutta tyypillisiä kuvantamisolosuhteissa tapahtuvia muutoksia simuloiden ja analyysimenetelmiä yhdistäen. Tutkimuksen kohteena ovat olleet erityisesti leikepaksuus ja kuvan signaali-kohinasuhde sekä tekstuurianalyysia täydentävän diffuusiotensorikuvauksen hyödyntäminen. Tässä luvussa arvioidaan tutkimustuloksia ja niiden merkitystä muun muassa lääketieteellisen käytännön näkökulmasta.

Tekstuurianalyysin tarkkuuteen, luotettavuuteen ja toistettavuuteen vaikuttavia tekijöitä on tämän työn lisäksi tarkasteltu muutamissa aiemmissä tutkimuksissa. Lyhyt kooste tässä työssä tutkituista ja eräistä muista tekijöistä on esitetty Taulukossa 9.

Taulukko 9. *Tekstuurianalyysin tarkkuuteen, luotettavuuteen ja toistettavuuteen vaikuttavia tekijöitä sekä niihin liittyviä tutkimuksia.*

Huomioitava tekijä	Aiempia tutkimuksia
Magneettikentän voimakkuus	Waugh ym. (2011) [84]
Kuvantamisprotokollat	Herlidou-Même ym. (2003) [87]; Mayerhoefer ym. (2009) [78]; Waugh ym. (2011) [84]
Leikepaksuus	Guggenbuhl ym. (2008) [81] (CT-kuvantaminen); Savio ym. (2010) [I]
Resoluutio	Jirák ym. (2004) [12]; Schad ja Lundervold (2006) [80]; Mayerhoefer ym. (2009) [82]
Signaali-kohinasuhde	Schad (2004) [94]; Waugh ym. (2011) [84]; Savio ym. (2011) [II]
Kontrasti	Mahmoud-Ghoneim ym. (2008) [90]
Intensiteetti ja sen vääristymät	Mahmoud-Ghoneim ym. (2008) [90]; Materka ja Strzelecki (2013) [91]
ROI-alueen asettelu ja koko	Harrison ym. (2010) [37]; Sikiö ym. (2015) [96]
Analyysidimensio (2D/3D)	Mahmoud-Ghoneim ym. (2003) [61]; Wagner ym. (2012) [59]
Täydentävät menetelmät	de Carvalho Alegro ym. (2012) [68]; Savio ym. (2015) [III]

7.1 Leikepaksuus

Luvussa 6.1 esitellyt, Osatyössä I [I] aiemmin julkaistut tulokset viittaavat siihen, että leikepaksuuden vaikutus MS-plakkien ja aivojen valkean aineen tekstuurianalyysiin on suhteellisen vähäinen. Lineaarisen ja epälineaarisen erotteluanalyysin tuottamat tulokset olivat lähes samankaltaisia. MS-plakkien ja normaalin tai normaalilta näyttävän valkean aineen keskimääräinen luokittelutarkkuus oli 92 %. Huomionarvoista on, että va-

paasti piirretyt ROI-alueet mahdollistivat vakiokokoisia ROI-alueita keskimäärin 4-6 prosenttiyksikköä paremmat luokittelutulokset. Valkean aineen ja normaalilta näyttävän valkean aineen luokittelutarkkuus oli sitä vastoin yleisesti varsin matala leikepaksuudesta, ROI-valinnasta tai luokittelualgoritmista riippumatta.

Tulosten yksityiskohtaisempi tarkastelu osoitti tekstuuriparametrien tilastollisen hajonnan eri leikepaksuuksien välillä suuremmaksi MS-plakkeja kuin valkean aineen alueita tarkasteltaessa [I]. Tämä voi viitata siihen, että leikepaksuuden valinta vaikuttaa voimakkaammin MS-plakkien kuin valkean aineen tekstuuriin. Varsinaisten luokittelutulosten ristiinvalidoinnin avulla arvioitu hajonta oli kuitenkin yleisesti ottaen vähäistä.

Edellä esitettyjen tutkimustulosten nojalla 1 mm leikkeiden korvaaminen 3 mm leikkeillä ei näyttäisi vaikuttavan olennaisesti MS-potilaista otettujen magneettikuvien tekstuuriin. Hyvin pieni leikepaksuus johtaisi kuitenkin matalaan signaali-kohinasuhteeseen ja heikentäisi siten tekstuurianalyysin luotettavuutta. Myös leikepaksuuden kasvaessa kudosten tekstuurien yksityiskohdat alkavat lopulta hävitä. Leikepaksuuden vaihdellessa 1 millimetrin ja 3 millimetrin välillä kumpaakaan edellä mainituista ilmiöistä ei havaittu. [I]

Tutkimustulosten perusteella useat parametrit voivat olla käyttökelpoisia MS-taudin tekstuurianalyysissa. Lisäksi ne ovat ainakin osittain riippumattomia magneettikuvien leikepaksuudesta (1 mm tai 3 mm). Vastaavaan tapaan Herlidou-Mêmen ym. fantomeja koskevat luokittelutulokset olivat likimain leikepaksuudesta riippumattomia [87]. On tosin syytä muistaa, että edellä mainitussa tutkimuksessa varsinkin testiaineiston koko oli melko pieni.

Magneettikuvan leikepaksuutta valittaessa on tarpeellista huomioida tarkasteltavien kudosten erityispiirteet. Harrison ym. totesivat MS-taudin tekstuuria käsitelleessä tutkimuksessaan, että normaalilta näyttävän harmaan aineen alueilta mitatut tekstuuriarvot vaihtelivat leikkeestä toiseen siirryttäessä enemmän kuin valkean aineen alueilta tai vaurio-alueilta mitatut [37]. Lisäksi on syytä huomioida Filippin ym. havainto leikepaksuuden valinnan merkityksestä MS-plakkien tilavuutta ja taudin etenemistä arvioitaessa [100].

7.2 Signaali-kohinasuhde

LDA- ja k-NN-menetelmien väliset erot olivat suhteellisen pieniä mitattaessa kohinan vaikutusta tekstuurianalyysin luokittelutarkkuuteen (Luku 6.2, Osatyö II [II]). Lähes kaikkien luokittelujen tarkkuus oli vähintään 90 %, kun aineiston signaali-kohinasuhde ylitti 20 dB. Leikepaksuustutkimuksen tavoin luokittelutulosten ristiinvalidoinnin avulla arvioitu hajonta oli vähäistä.

Parhaat tulokset saatiin rasvan ja lihaskudosten (suora reisilihas tai ulompi reisilihas) välisissä luokitteluissa. Kyseiset luokittelut olivat lähes virheettömiä, jos signaali-kohinasuhde oli yli 30 dB. Signaali-kohinasuhteen pudotessa kohti 15 desibeliä heikkeni luokittelutarkkuus jyrkästi kohti 50 prosenttia. Myös luuydin erottui tutkimustulosten perusteella lihaskudoksesta pääsääntöisesti hyvin, kun lihaskudoksen signaalikohinasuhde oli vähintään 25 dB ja luuytimen vastaavasti vähintään 20 dB. Signaali-kohinasuhteen pudotessa luokittelutarkkuus kuitenkin heikkeni jyrkemmin kuin rasvaa lihaskudoksiin vertaessa. [II]

Signaali-kohinasuhteen pieneneminen vaikutti rasvan ja luuytimen keskinäiseen vertailuun muita tarkasteluja vähemmän. Viidestä muusta tarkasteltavasta kudosparista poiketen nämä kudostyypit erottuivat toisistaan vähintään 70-prosenttisesti vielä toiseksi alimmassa SNR-luokassa. Selvästi vaikeimmaksi tehtäväksi osoittautui kahden eri lihaskudoksen luokittelu. Yli 80 prosentin luokittelutarkkuuteen päästiin vain kahdessa ylimässä SNR-luokassa. [II] Skochin ym. suorittamassa terveiden ja sairaiden lihasten tekstuurianalyysissä saavutettiin vastaavaan tapaan noin 80 prosentin luokittelutarkkuus [101]. Heidän tutkimuksessaan terveiden ja sairaiden lihasten erot eivät olleet nähtävissä silmämääräisesti.

Tekstuuriparametrit ovat Schadin ja Lundervoldin toteamalla tavalla alttiita kohinan vaikutukselle [80]. Heidän mukaansa kudosten todellinen tekstuuri tulee esiin kohinan pudotessa tietyn rajan alapuolelle, minkä jälkeen signaali-kohinasuhteen parantumisella ei ole suurta vaikutusta tekstuuriparametrien arvoihin. Samankaltainen havainto tehtiin Osatyössä II [II], jossa signaali-kohinasuhteen tyypillinen, tarkasteltavasta kudoksesta riippuva kynnysarvo oli noin 20-30 dB.

Reiden magneettikuviin sovelletun tekstuurianalyysitutkimuksen perusteella on vaikeaa määrittää yksittäistä signaali-kohinasuhteen kynnysarvoa, jonka alapuolella luokittelu muuttuu liian epätarkaksi. Edellä mainittu kynnysarvo on riippuvainen tarkasteltavista kudoksista ja käytetyistä tekstuuriparametreista. Kynnysarvon ylittyessä oli kudosten luokittelutarkkuus useimmissa tapauksissa vähintään 90 %. Luokittelutarkkuus ei tällöin riippunut merkittävästi yksittäisen algoritmin (lineaarinen erotteluanalyysi tai k lähintä naapuria) valinnasta. Lihasyhmien keskinäinen luokittelu edellyttäisi kuitenkin selvästi korkeampaa signaali-kohinasuhdetta. [II]

Jos tekstuurianalyysin opetusaineistoon sisältyvissä kuvissa on kohtuullinen määrä kohinaa, saattaa tekstuurianalyysin tarkkuus jopa parantua verrattuna tilanteeseen, jossa opetusaineisto sisältää hyvin vähän häiriölähteitä [II]. Saman havainnon ovat tehneet aiemmin muun muassa Mayerhoefer ym. [78]. Kyseinen havainto saattaa johtua siitä,

että algoritmin on otettava opetustilanteessa laajasti huomioon tavanomaisessa aineistossa väistämättä esiintyvä satunnaisuus, eikä sitä pystytä ylisovittamaan käytössä olevaan aineistoon.

7.3 Tekstuurianalyysin ja diffuusiotensorikuvauksen yhdistäminen

PPMS- ja RRMS-potilaiden magneettikuvista löydettiin tilastollisesti merkitseviä puolieroja sekä tekstuurianalyysillä että diffuusiotensorimittauksilla (Luku 6.3, Osatyö III [III]). Erityisesti thalamuksen alueella mitatut puolierot olivat RRMS-potilailla selvästi suurempia kuin PPMS-potilailla.

Tekstuurianalyysissä vasemman ja oikean aivopuoliskon välisiä puolieroja löydettiin eniten RRMS-potilaiden thalamuksesta, jossa 23 % tekstuuriparametreista osoitti tilastollisesti merkitseviä ($p < 0,01$) eroja vasemman ja oikean aivopuoliskon välillä. Myös DTI-analyysissä RRMS-potilaiden aivoissa havaittiin suurempia puolieroja kuin PPMS-potilailla. Tilastollisesti merkitseviä puolieroja oli selvästi eniten RRMS-potilaiden caudate nucleuksen alueella, lisäksi puolieroja havaittiin centrum semiovalen ja thalamuksen alueilla. RRMS-potilailta mitatut puolierot korostuivat PPMS-potilaiden puolieroihin verrattuna varsinkin aivojen harmaan aineen alueella. [III]

PPMS- ja RRMS-tautityyppien luokittelussa päästiin noin 80 prosentin luokittelutarkkuuteen. Luokittelumenetelmistä lähimmän naapurin menetelmä tuotti parhaat tulokset, joskin NDA-luokittelualgoritmi oli lähes yhtä tarkka. Edellä mainituissa tapauksissa ei havaittu suuria eroja vain kahta tekstuuriparametria hyödyntävän luokittelijan ja sekä kahta tekstuuriparametria että kahta diffuusiotensoriparametria käyttävän luokittelijan välillä. Sitä vastoin LDA-menetelmän luokittelutulokset olivat selvästi epälineaaristen tai diskreettien algoritmien antamia tuloksia heikompia etenkin pelkkiä tekstuuriparametreja käytettäessä. [III]

Tulokset viittaavat siihen, että tekstuurianalyysi voi helpottaa MS-tautityypin selvittämistä. Tekstuurianalyysin ja DTI-mittaustulosten samanaikainen hyödyntäminen ei parantanut luokittelutuloksia, vaikka tekstuuriparametrien ja FA- tai ADC-arvojen välillä ei löydettykään tilastollisesti merkitseviä korrelaatioita. DTI-menetelmä näyttäisi kuitenkin auttavan taudin etenemisestä kertovien puolierojen löytämisessä. Onkin mahdollista, että sekä tekstuurianalyysi että DTI-menetelmä tarjoavat käytännön potilastyön kannalta hyödyllistä lisätietoa, vaikka niiden samanaikainen käyttö ei lisäisikään MS-tautityypin diagnosointitarkkuutta. [III]

Aiemmin on havaittu MS-taudin voivan alkaa kehittyä harmaalla alueella ja edetä myöhemmin valkeaan aineeseen [102-103]. Riccitelli ym. havaitsivat kognitiivisista vaikeuksista kärsivien RRMS-potilaiden vasemmassa thalamuksessa merkittäviä muutoksia T2-painotteisissa magneettikuvissa [104]. Myös Minagar ym. ovat kehottaneet kiinnittämään erityistä huomiota thalamuksessa tapahtuviin muutoksiin [57]. Fabiano ym. puolestaan löysivät ADC-mittauksissa puolieroja RRMS-potilaiden vasemman ja oikean aivopuoliskon välillä [105]. Nämä havainnot ovat sopusoinnussa Osatyössä III [III] esitettyjen tulosten kanssa.

7.4 Muiden tekijöiden vaikutus tekstuurianalyysin tarkkuuteen

Tekstuurianalyysin tarkkuuteen vaikuttavien tekijöiden joukko on varsin laaja. Periaatteessa ainakin kaikki edellä mainitut tekijät tulisi ottaa huomioon tekstuurianalyysin käyttöä suunniteltaessa. Tuloksiin saattavat lisäksi vaikuttaa monet vielä toistaiseksi tuntemattomat seikat.

Schad on korostanut fantomien merkitystä tekstuurianalyysin luotettavuuden arvioinnissa [94] Zhang ym. ovat todenneet jatkotutkimusta tarvittavan, jotta tekstuurianalyysi voisi vakiinnuttaa asemansa vakaana, päivittäisessä potilastyössä käytettävänä menetelmänä. Mahdollisia muutoksia tekstuuriin voi aiheutua esimerkiksi RF-kentän epähomogeenisuuksista, mikä edellyttää jo kuvantamisvaiheessa tarkkaa sekvenssien ja kuvantamisparametrien valintaa. [8]

Vaikka monet tekniset haasteet ovat ainakin osittain ratkaistavissa [106-108], on potilaasta johtuvien kuvavääristymien korjaaminen erityisen hankalaa. Joissakin tutkimuksissa [35][109] on todettu, että potilaan liikehännän aiheuttamien kuvavääristymien vaikutus tekstuuriin on hankalasti arvioitavissa, koska kyseisiä vääristymiä on vaikea erottaa muusta kuvadatasta.

Yksittäisten tekijöiden osaoptimointi saattaa johtaa muiden virhelähteiden vaikutuksen kasvuun. Esimerkiksi kuvantamisajan lyhentämiseen tähtäävä, Fourier-muunnoksessa käytettävien vaihekoodausten (phase-encoding lines) lukumäärän pienentäminen voi signaali-kohinasuhteen heikentämisen lisäksi muuttaa kuvan sisältämän kohinajakauman muotoa ja siten sen tekstuuria [110]. Matalaa signaali-kohinasuhdetta voidaan yrittää parantaa muun muassa kasvattamalla kuvanmuodostuskertojen (acquisitions) määrää tai pidentämällä TR-aikaa. Kokonaiskuvantamisajan kasvaessa potilaan liikehännästä johtuvien kuvavääristymien riski lisääntyy, eikä tekstuurianalyysin tarkkuutta ja tulosten luotettavuutta pystytä kaikissa olosuhteissa takaamaan.

Käytännön kuvantamistilanteissa voidaan tähdätä kompromissiin esimerkiksi signaali-kohinasuhteen ja kuvausajan välillä, ellei korkea signaali-kohinasuhde ole tutkimuksen onnistumisen ja tulosten tulkinnan välttämätön edellytys. Vastaavasti korkea resoluutio ja ohut leikepaksuus vaikeuttavat hyvän signaali-kohinasuhteen saavuttamista, mikä hankaloittaa varsinkin pienten kudosalueiden analysointia.

Tekstuurianalyysin tulisi olla mahdollisimman luotettavaa ja sen tuottamien tulosten toistettavia. Erityisesti ROI-alueiden määrittely on hankalasti dokumentoitava ja toistettava vaihe. Harrison ym. ovat huomauttaneet, että verrattaessa MS-plakkien kohdalle vapaasti piirrettyjen ROI-alueiden tekstuuria vakiokokoisten, suorakaiteen muotoisten ROI-alueiden tekstuuriin tulee varmistaa, etteivät ROI-alueet sisällä osia viereisistä kudoksista. Erot MS-plakkien muodossa ja koossa voivat johtaa osittaisilavuusvaikutuksen kasvuun myös vakiokokoisilla ROI-alueilla. Tällä saattaa olla merkittävä vaikutus tekstuurianalyysin tarkkuuteen. [37]

Osatyössä I [I] toteutetun MS-plakkien ja valkean aineen luokittelun tarkkuus riippui jossain määrin tarkastelualueen koosta ja muodosta. Kuten edellä todettiin, vapaasti piirretty ROI-alueet mahdollistivat muutamaa prosenttiyksikköä paremmat luokittelutulokset vakiokokoisten ROI-alueiden käyttöön verrattuna. Harrison ym. ovat toisaalta korostaneet vakiokokoisten ja -muotoisten ROI-alueiden myönteisiä vaikutuksia tekstuurianalyysiprosessiin ja sen tuloksiin [37]. Niitä ovat muun muassa tulosten riippumattomuus leikevalinnasta, asettelun nopeus ja analyysin toistettavuus. He myös muistuttivat, että vapaasti piirrettävien ROI-alueiden liian pieni koko saattaa heikentää tilastollisen analyysin luotettavuutta. Erityisesti Osatyössä I [I] käytettyjen ROI-alueiden koko on vaihdellut merkittävästi, mikä on otettava huomioon tuloksia arvioitaessa.

Sikiö ym. ovat tutkineet ROI-alueen koon vaikutusta tekstuurianalyysin tarkkuuteen ja havainneet, että se tulisi sovittaa tarkasteltavan kudoksen rakenteeseen [96]. Heidän tutkimustulostensa perusteella sekä liian suurten että liian pienten ROI-alueiden käyttö voi hävittää tutkittavan tekstuurin. Larroza ym. ovat arvioineet automatisoituun ROI-valintaan liittyvää lisätutkimusta tarvittavan tekstuurianalyysin luotettavuuden parantamiseksi [1].

Myös tekstuurianalyysin dimension valinta edellyttää kokonaisvaltaista harkintaa. Vaikka kolmiulotteinen tekstuurianalyysi tuottaa useiden tutkimustulosten [59-63] valossa kaksiulotteista vastinpariaan muutamaa prosenttiyksikköä parempia luokittelutuloksia, se voi olla kaksiulotteista menetelmää alttiimpi häiriötekijöille. Arkisen potilastyön kannalta on aiheellista puntaroida tarkoin ajankäytön ja analyysitarkkuuden paranemisen suhdetta.

Kolmiulotteinen tekstuurianalyysi vaatii yleensä huomattavasti kaksiulotteista menetelmää pidemmän toteutusajan. Lisäksi potilaskuvat ovat toisinaan saatavilla vain kaksiulotteisessa muodossa. Ellei luokittelumenetelmien tarkkuudessa ole merkittävää eroa, onkin kaksiulotteinen tekstuurianalyysi yleensä käytännössä suositeltavampi menetelmä.

8. JOHTOPÄÄTÖKSET

Tässä työssä tutkittiin magneettikuviin sovelletun tekstuurianalyysin tarkkuuteen ja luotettavuuteen vaikuttavia tekijöitä. Erityisesti tarkasteltiin kuvien leikepaksuutta ja signaali-kohinasuhdetta sekä tekstuurianalyysia täydentävän diffuusiotensorikuvausdatan hyödyntämistä. Aineistona käytettiin MS-potilaiden aivoista ja terveiden kohdehenkilöiden reiden alueelta otettuja magneettikuvia. Tutkimustulosten yleistettävyyttä pyrittiin parantamaan ristiinvalidoinnilla erillisiä opetus- ja testiaineistoja hyödyntäen.

Tutkimustulosten perusteella tekstuurianalyysi on suhteellisen vakaa menetelmä. Pienet muutokset leikepaksuudessa tai signaali-kohinasuhteessa vaikuttivat vain vähäisessä määrin tekstuurianalyysin luokittelutarkkuuteen. Ristiinvalidoinnin avulla arvioitu luokittelutulosten hajonta oli molemmissa tapauksissa vähäistä. Diffuusiotensorikuvantamisen soveltaminen yhdessä tekstuurianalyysin kanssa ei kasvattanut MS-tautityyppiluokittelun tarkkuutta, mutta paljasti taudin etenemiseen viittaavia puolieroja.

Esitellyt havainnot ovat lupaavia ja viittaavat siihen, että tekstuurianalyysillä on edellytyksiä yleistyä ja antaa luotettavia tuloksia kliinisessä hoitotyössä. Tekstuurianalyysin vakiintuminen osaksi päivittäistä hoitorutiinia edellyttää silti edelleen jatkotutkimusta sen tarkkuuteen ja tulosten toistettavuuteen vaikuttavista tekijöistä kaikissa tekstuurianalyysin vaiheissa. Lisäksi erityistä huomiota tulee kiinnittää tällaisten tekijöiden vaikutuksen kasautumiseen varsinkin silloin, kun ne ovat signaali-kohinasuhteen ja leikepaksuuden tai resoluution tavoin toisistaan riippuvia.

LÄHDELUETTELO

- [1] Larroza A, Bodí V, Moratal D. Texture analysis in magnetic resonance imaging: review and considerations for future applications. Teoksessa Constantinides C (toim.). Assessment of Cellular and Organ Function and Dysfunction Using Direct and Derived MRI Methodologies. InTechOpen 2016; 75-106.
- [2] Hájek M, Dezortova M, Materka A, Lerski R (toim.). Texture analysis for magnetic resonance imaging. Med4publishing, Prague 2006. 234 sivua.
- [3] Materka A, Strzelecki M. Texture analysis methods – a review. COST B11 – technical report. Technical University of Lodz 1998.
- [4] Zhang J, Tong L, Wang L, Li N. Texture analysis of multiple sclerosis: a comparative study. Journal of Magnetic Resonance Imaging 2008; 26: 1160-1166.
- [5] Haralick RM. Statistical and structural approaches to texture. Proc IEEE, 1979; 67: 786-804.
- [6] Kassner A, Thornhill RE. Texture analysis: a review of neurologic MR imaging applications. American Journal of Neuroradiology 2010; 31: 809-816.
- [7] Castellano G, Bonilha L, Li LM, Cendes F. Texture analysis of medical images. Clinical Radiology 2004; 59: 1061-1069.
- [8] Zhang Y. MRI texture analysis in multiple sclerosis. International Journal of Biomedical Imaging 2012; 2012: 762804.
- [9] Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on featured distributions. Pattern Recognition 1996; 29: 51-59.
- [10] Feng Z, Rong P, Cao P, Zhou Q, Zhu W, Yan Z, Liu Q, Wang W. Machine learning-based quantitative texture analysis of CT images of small renal masses: Differentiation of angiomyolipoma without visible fat from renal cell carcinoma. European Radiology 2018; 28: 1625-1633.
- [11] Lerski RA, de Certaines JD, Duda D, Klonowski W, Yang G, Coatrieux JL, Azabou N, Eliat PA. Application of texture analysis to muscle MRI: 2 – technical recommendations. EPJ Nonlinear Biomedical Physics 2015; 3: 2.
- [12] Jiráček D, Dezortová M, Hájek M. Phantoms for texture analysis of MR images. Long-term and multi-center study. Medical Physics 2004; 31: 616-622.
- [13] Guyon I, Elisseeff A. An introduction to variable and feature selection. Journal of Machine Learning Research 2003; 3: 1157-1182.
- [14] Depeursinge A, Foncubierta-Rodriguez A, Van De Ville D, Müller H. Three-dimensional solid texture analysis in biomedical imaging: review and opportunities. Medical Image Analysis 2014; 18: 176-196.

- [15] Antel SB, Collins DL, Bernasconi N, Andermann F, Shinghal R, Kearney RE, Arnold DL, Bernasconi A. Automated detection of focal cortical dysplasia lesions using computational models of their MRI characteristics and texture analysis. *Neuroimage* 2003; 19: 1748-1759.
- [16] Hastie T, Tibshirani R, Friedman J. The elements of statistical learning. Springer series in statistics. Springer, New York 2001.
- [17] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2002; 24: 509-522.
- [18] Zacharaki EI, Wang S, Chawla S, Soo Yoo D, Wolf R, Melhem ER, Davatzikos C. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magnetic Resonance in Medicine* 2009; 62: 1609–1618.
- [19] Rathore S, Iftikhar MA, Hussain M, Jalil A. Texture analysis for liver segmentation and classification: a survey. *Frontiers of Information Technology*, 2011; 121-126.
- [20] Campadelli P, Casiraghi E, Esposito A. Liver segmentation from computed tomography scans: a survey and a new algorithm. *Artificial Intelligence in Medicine* 2009; 45: 185-196.
- [21] Sharma N, Aggarwal LM. Automated medical image segmentation techniques. *Journal of Medical Physics* 2010; 35: 3–14.
- [22] Chang KI, Bowyer KW, Sivagurunath M. Evaluation of texture segmentation algorithms. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1999; 1: 294-299.
- [23] Pham DL, Xu C, Prince JL. Current methods in medical image segmentation. *Annual Review of Biomedical Engineering* 2000; 2: 315-337.
- [24] Muneeswaran K, Ganesan L, Arumugam S, Soundar KR. Texture image segmentation using combined features from spatial and spectral distribution. *Pattern Recognition Letters* 2006; 27: 755–764.
- [25] Kekre HB, Gharge S. Texture based segmentation using statistical properties for mammographic images. *International Journal of Advanced Computer Science Appl* 2010; 1: 102-107.
- [26] Tesar L, Smutek D, Shimizu A, Kobatake H. Medical image segmentation using cooccurrence matrix based texture features calculated on weighted region. *Proceedings of the third conference on IASTED International Conference: Advances in Computer Science and Technology* 2007; 243-248.
- [27] Kruggel F, Paul JS, Gertz HJ. Texture-based segmentation of diffuse lesions of the brain's white matter. *Neuroimage* 2008; 39: 987-996.
- [28] Martinez GV. Introduction to MRI physics. *Methods in Molecular Biology* 2018; 1718: 3-19.
- [29] Bloch F. Nuclear induction. *Physical Review* 1946; 70: 460-474.

- [30] Liang ZP, Lauterbur PC. Principles on magnetic resonance imaging: A signal processing perspective. IEEE Press, New York 2000. 416 sivua.
- [31] Fetit AE, Novak J, Peet AC, Arvanitits TN. Three-dimensional textural features of conventional MRI improve diagnostic classification of childhood brain tumours. *NMR in Biomedicine* 2015; 28: 1174-1184.
- [32] Yu O, Mauss Y, Namer IJ, Chambron J. Existence of contralateral abnormalities revealed by texture analysis in unilateral intractable hippocampal epilepsy. *Journal of Magnetic Resonance Imaging* 2001; 19: 1305-1310.
- [33] Suoranta S, Holli-Helenius K, Koskenkorva P, Niskanen E, Könönen M, Äikiä M, Eskola H, Kälviäinen R, Vanninen R. 3D texture analysis reveals imperceptible MRI textural alterations in the thalamus and putamen in progressive myoclonic epilepsy type 1, EPM1. *PLoS One* 2013; 8: e69905.
- [34] Maani R, Yang YH, Kalra S. Voxel-based texture analysis of the brain. *PLoS One* 2015; 10: e0117759.
- [35] Freeborough PA, Fox NC. MR image texture analysis applied to the diagnosis and tracking of Alzheimer's disease. *IEEE Transactions on Medical Imaging* 1998; 17: 475-479.
- [36] Sikiö M, Holli-Helenius KK, Harrison LC, Ryymin P, Ruottinen H, Saunamäki T, Eskola HJ, Elovaara I, Dastidar P. MR image texture in Parkinson's disease: a longitudinal study. *Acta Radiologica* 2015; 56: 97-104.
- [37] Harrison LC, Raunio M, Holli KK, Luukkaala T, Savio S, Elovaara I, Soimakallio S, Eskola HJ, Dastidar P. MRI texture analysis in multiple sclerosis: toward a clinical analysis protocol. *Academic Radiology* 2010; 17: 696-707.
- [38] Ko ES, Kim JH, Lim Y, Han BK, Cho EY, Nam SJ. Assessment of invasive breast cancer heterogeneity using whole-tumor magnetic resonance imaging texture analysis: Correlations with detailed pathological findings. *Medicine* 2016; 95: e2453.
- [39] Li Z, Mao Y, Li H, Yu G, Wan H, Li B. Differentiating brain metastases from different pathological types of lung cancers using texture analysis of T1 postcontrast MR. *Magnetic Resonance in Medicine* 2016; 76: 1410-1419.
- [40] Wibmer A, Hricak H, Gondo T, Matsumoto K, Veeraraghavan H, Fehr D, Zheng J, Goldman D, Moskowitz C, Fine SW, Reuter VE, Eastham J, Sala E, Vargas HA. Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores. *European Radiology* 2015; 25: 2840-2850.
- [41] Fox MJ, Gibbs P, Pickles MD. Minkowski functionals: An MRI texture analysis tool for determination of the aggressiveness of breast cancer. *Journal of Magnetic Resonance Imaging* 2016; 43: 903-910.
- [42] Harrison LC, Luukkaala T, Pertovaara H, Saarinen TO, Heinonen TT, Jarvenpää R, Soimakallio S, Kellokumpu-Lehtinen PL, Eskola HJ, Dastidar P. Non-Hodgkin lymphoma response evaluation with MRI texture classification. *Journal of Experimental & Clinical Cancer Research* 2009; 28: 87.

- [43] Holli KK, Wäljas M, Harrison L, Liimatainen S, Luukkaala T, Ryymin P, Eskola H, Soimakallio S, Öhman J, Dastidar P. Mild traumatic brain injury: tissue texture analysis correlated to neuropsychological and DTI findings. *Academic Radiology* 2010; 17: 1096-1102.
- [44] Zhang X, Gao X, Liu BJ, Ma K, Yan W, Liling L, Yuhong H, Fujita H. Effective staging of fibrosis by the selected texture features of liver: Which one is better, CT or MR imaging? *Computerized Medical Imaging and Graphics* 2015; 46: 227-236.
- [45] Herlidou S, Rolland Y, Bansard JY, Le Rumeur E, de Certaines JE. Comparison of automated and visual texture analysis in MRI: characterization of normal and diseased skeletal muscle. *Journal of Magnetic Resonance Imaging* 1999; 17: 1393-1397.
- [46] Nakai R, Azuma T, Sudo M, Urayama S, Takizawa O, Tsutsumi S. MRI analysis of structural changes in skeletal muscles and surrounding tissues following long-term walking exercise with training equipment. *Journal of Applied Physiology* 2008; 105: 958-963.
- [47] Mahmoud-Ghoneim D, Cherel Y, Lemaire L, de Certaines JD, Maniere A. Texture analysis of magnetic resonance images of rat muscles during atrophy and regeneration. *Journal of Magnetic Resonance Imaging* 2006; 24: 167-171.
- [48] Nketiah G, Savio S, Dastidar P, Nikander R, Eskola H, Sievänen H. Detection of exercise load-associated differences in hip muscles by texture analysis. *Scandinavian Journal of Medicine & Science in Sports* 2015; 25: 428-434.
- [49] Lespessailles E, Gadois C, Kousignian I, Neveu JP, Fardellone P, Kolta S, Roux C, Do-Huu JP, Benhamou CL. Clinical interest of bone texture analysis in osteoporosis: a case control multicenter study. *Osteoporosis International* 2008; 19: 1019-1028.
- [50] Lerski RA, Straughan K, Schad LR, Boyce D, Blüml S, Zuna I. MR image texture analysis - An approach to tissue characterization. *Journal of Magnetic Resonance Imaging* 1993; 11: 873-887.
- [51] Schad LR, Blüml S, Zuna I. MR tissue characterization of intracranial tumors by means of texture analysis. *Journal of Magnetic Resonance Imaging* 1993; 11: 889-896.
- [52] McQualter JL, Bernard CC. Multiple sclerosis: a battle between destruction and repair. *Journal of Neurochemistry* 2007; 100: 295-306.
- [53] Ramagopalan SV, Dobson R, Meier UC, Giovannoni G. Multiple sclerosis: risk factors, prodromes, and potential causal pathways. *The Lancet Neurology* 2010; 9: 727-739.
- [54] Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, Correale J, Fazekas F, Filippi M, Freedman MS, Fujihara K, Galetta SL, Hartung HP, Kappos L, Lublin FD, Marrie RA, Miller AE, Miller DH, Montalban X, Mowry EM, Sorensen PS, Tintoré M, Traboulsee AL, Trojano M, Uitdehaag BMJ, Vukusic S, Waubant E, Weinshenker BG, Reingold SC, Cohen JA. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *The Lancet Neurology* 2018; 17: 162-173.

- [55] Trojano M, Paolicelli D, Bellacosa A, Cataldo S. The transition from relapsing-remitting MS to irreversible disability: clinical evaluation. *Neurological Sciences* 2003; 24 (Suppl 5): S268-270.
- [56] Antel J, Antel S, Caramanos Z, Arnold DL, Kuhlmann T. Primary progressive multiple sclerosis: part of the MS disease spectrum or separate entity? *Acta Neuropathologica* 2012; 123: 627-638.
- [57] Minagar A, Barnett MH, Benedict RH, Pelletier D, Pirko I, Sahraian MA, Frohman E, Zivadinov R. The thalamus and multiple sclerosis: modern views on pathologic, imaging, and clinical aspects. *Neurology* 2013; 80: 210-219.
- [58] Bento M, Rittner L, Appenzeller S, Lapa A, Lotufo R. Analysis of brain white matter hyperintensities using pattern recognition techniques. Ourselin S, Haynor DR (toim.). *Medical Imaging 2013: Image Processing, Proc. of SPIE Vol. 8669*, 86693P.
- [59] Wagner F, Gryanik A, Schulz-Wendtland R, Fasching PA, Wittenberg T. 3D characterization of texture: Evaluation for the potential application in mammographic mass diagnosis. *Biomedizinische Technik/Biomedical Engineering* 2012; 57 (Suppl 1): 490-493.
- [60] Georgiadis P, Cavouras D, Kalatzis I, Glotsos D, Athanasiadis E, Kostopoulos S, Sifaki K, Malamas M, Nikiforidis G, Solomou E. Enhancing the discrimination accuracy between metastases, gliomas and meningiomas on brain MRI by volumetric textural features and ensemble pattern recognition methods. *Journal of Magnetic Resonance Imaging* 2009; 27: 120-130.
- [61] Mahmoud-Ghoneim D, Toussaint G, Constans JM, de Certaines JD. Three dimensional texture analysis in MRI: a preliminary evaluation in gliomas. *Journal of Magnetic Resonance Imaging* 2003; 21: 983-987.
- [62] Chen W, Giger ML, Li H, Bick U, Newstead GM. Volumetric texture analysis of breast lesions on contrast-enhanced magnetic resonance images. *Magnetic Resonance in Medicine* 2007; 58: 562-571.
- [63] Nketiah G. Comparison of 2D and 3D texture analyses of functionally different hip muscles. <http://URN.fi/URN:NBN:fi:ty-201303181092>
- [64] Baliyan V, Das CJ, Sharma R, Gupta AK. Diffusion weighted imaging: Technique and applications. *World Journal of Radiology* 2016; 8: 785-798.
- [65] Hakulinen U, Brander A, Ryymin P, Öhman J, Soimakallio S, Helminen M, Dastidar P, Eskola H. Repeatability and variation of region-of-interest methods using quantitative diffusion tensor MR imaging of the brain. *BMC Medical Imaging* 2012; 12: 30.
- [66] Kolasa M, Hakulinen U, Brander A, Hagman S, Dastidar P, Elovaara I, Sumelahti ML. Diffusion tensor imaging and disability progression in multiple sclerosis: A 4-year follow-up study. *Brain and Behavior* 2019; 9: e01194.
- [67] Eliat PA, Olivié D, Saïkali S, Carsin B, Saint-Jalmes H, de Certaines JD. Can dynamic contrast-enhanced magnetic resonance imaging combined with texture analysis differentiate malignant glioneuronal tumors from other glioblastoma? *Neurology Research International* 2012; 2012: 195176.

- [68] de Carvalho Alegro M, Valotta Silva A, Yumi Bando S, de Deus Lopes R, Martins de Castro LH, Hungtsu W, Moreira-Filho CA, Amaro E Jr. Texture analysis of high resolution MRI allows discrimination between febrile and afebrile initial precipitating injury in mesial temporal sclerosis. *Magnetic Resonance in Medicine* 2012; 68: 1647-1653.
- [69] de Certaines JD, Larcher T, Duda D, Azzabou N, Eliat PA, Escudero LM, Pinheiro AMG, Yang G, Coatrieux JL, Snezhko E, Shukelovich A, Pereira M, Lerski RA. Application of texture analysis to muscle MRI: 1 – What kind of information should be expected from texture analysis? *EPJ Nonlinear Biomedical Physics* 2015; 3: 3.
- [70] Puig D, Garcia MA. Automatic texture feature selection for image pixel classification. *Pattern Recognition*, 2006; 39: 1996-2009.
- [71] Pusey E, Lufkin RB, Brown RK, Solomon MA, Stark DD, Tarr RW, Hanafee WN. Magnetic resonance imaging artifacts: mechanism and clinical significance. *Radiographics* 1986; 6: 891-911.
- [72] Dougherty G. Digital image processing for medical applications. Cambridge University Press 2009. 462 sivua.
- [73] Weishaupt D, Köchli VD, Marincek B. How does MRI work?: An introduction to the physics and function of magnetic resonance imaging. Springer-Verlag, Berlin Heidelberg 2006.
- [74] Link TM, Majumdar S, Peterfy C, Daldrup HE, Uffmann M, Dowling C, Steinbach L, Genant HK. High resolution MRI of small joints: impact of spatial resolution on diagnostic performance and SNR. *Journal of Magnetic Resonance Imaging* 1998; 16: 147-155.
- [75] Macovski A. Noise in MRI. *Magnetic Resonance in Medicine* 1996; 36: 494-497.
- [76] Parker DL, Gullberg GT. Signal-to-noise efficiency in magnetic resonance imaging. *Medical Physics* 1990; 17: 250-257.
- [77] Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine* 1995; 34: 910-914.
- [78] Mayerhoefer ME, Szomolanyi P, Jirák D, Materka A, Trattnig S. Effects of MRI acquisition parameter variations and protocol heterogeneity on the results of texture analysis and pattern discrimination: an application-oriented study. *Medical Physics* 2009; 36: 1236-1243.
- [79] Kovalev VA, Petrou M, Bondar YS. Texture anisotropy in 3-D images. *IEEE Transactions on Image Processing* 1999; 8: 346-360.
- [80] Schad LR, Lundervold A. Influence of resolution and signal to noise ratio on MR image texture. Teoksessa Hájek M, Dezortova M, Materka A, Lerski R (toim.). *Texture Analysis for Magnetic Resonance Imaging*. Med4publishing, Prague 2006; 129–149.
- [81] Guggenbuhl P, Chappard D, Garreau M, Bansard JY, Chales G, Rolland Y. Reproducibility of CT-based bone texture parameters of cancellous calf bone samples: influence of slice thickness. *European Journal of Radiology* 2008; 67: 514-520.

- [82] Mayerhoefer ME, Szomolanyi P, Jirák D, Berg A, Materka A, Dirisamer A, Trattnig S. Effects of magnetic resonance image interpolation on the results of texture-based pattern classification: a phantom study. *Investigative Radiology* 2009; 44: 405-411.
- [83] Brynolfsson P, Nilsson D, Torheim T, Asklund T, Karlsson CT, Trygg J, Nyholm T, Garpebring A. Haralick texture features from apparent diffusion coefficient (ADC) MRI images depend on imaging and pre-processing parameters. *Scientific Reports* 2017; 7: 4041.
- [84] Waugh SA, Lerski RA, Bidaut L, Thompson AL. The influence of field strength and different clinical breast MRI protocols on the outcome of texture analysis using foam phantoms. *Medical Physics* 2011; 38: 5058-5066.
- [85] Ford J, Dogan N, Young L, Yang F. Quantitative radiomics: Impact of pulse sequence parameter selection on MRI-based textural features of the brain. *Contrast Media & Molecular Imaging* 2018: 1729071.
- [86] Mayerhoefer ME, Schima W, Trattnig S, Pinker K, Berger-Kulemann V, Ba-Ssalamah A. Texture-based classification of focal liver lesions on MRI at 3.0 Tesla: a feasibility study in cysts and hemangiomas. *Journal of Magnetic Resonance Imaging* 2010; 32: 352-359.
- [87] Herlidou-Même S, Constans JM, Carsin B, Olivie D, Eliat PA, Nadal-Desbarats L, Gondry C, Le Rumeur E, Idy-Peretti I, de Certaines JD. MRI texture analysis on texture test objects, normal brain and intracranial tumors. *Journal of Magnetic Resonance Imaging* 2003; 21: 989-993.
- [88] Fruehwald-Pallamar J, Hesselink JR, Mafee MF, Holzer-Fruehwald L, Czerny C, Mayerhoefer ME. Texture-based analysis of 100 MR examinations of head and neck tumors – is it possible to discriminate between benign and malignant masses in a multicenter trial? *Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren* 2016; 188: 195–202.
- [89] Xiao YD, Paudel R, Liu J, Ma C, Zhang ZS, Zhou SK. MRI contrast agents: Classification and application (Review). *International Journal of Molecular Medicine* 2016; 38: 1319-1326.
- [90] Mahmoud-Ghoneim D, Alkaabi MK, de Certaines JD, Goettsche FM. The impact of image dynamic range on texture classification of brain white matter. *BMC Medical Imaging* 2008; 8: 18.
- [91] Materka A, Strzelecki M. On the importance of MRI nonuniformity correction for texture analysis. *IEEE International Conference on Signal Processing, Algorithms, Architectures, Arrangements and Applications* 2013; 118-123.
- [92] Styner M, Van Leemput K. Retrospective evaluation and correction of intensity inhomogeneities. Teoksessa Landini L, Positano V, Santarelli MF (toim.). *Advanced Image Processing in Magnetic Resonance Imaging*. Boca Raton: Taylor & Francis CRC Press 2005; 145-168.
- [93] Materka A, Strzelecki M, Lerski R, Schad L. Feature evaluation of texture test objects for magnetic resonance imaging. Teoksessa Pietikäinen M (toim.). *Texture Analysis in Machine Vision. Series in Machine Perception and Artificial Intelligence*. World Scientific, Singapore 2000; 40: 197-206.

- [94] Schad LR. Problems in texture analysis with magnetic resonance imaging. *Dialogues in Clinical Neuroscience* 2004; 6: 235-242.
- [95] Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Journal of Magnetic Resonance Imaging* 2004; 22: 81–91.
- [96] Sikiö M, Holli-Helenius KK, Harrison L, Ryymin P, Dastidar P, Eskola H. The effect of region of interest size on textural parameters. *IEEE Proceedings of Image and Signal Processing and Analysis, Zagreb, Croatia September 7-9 2015*; 149-153.
- [97] Strzelecki M, Szczypinski P, Materka A, Klepaczko A. A software tool for automatic classification and segmentation of 2D/3D medical images. *Nuclear Instruments and Methods in Physics Research Section A* 2013; 702: 137-140.
- [98] Ahmad MS, Shadid Naweed M, Nisa M. Application of texture analysis in the assessment of chest radiographs. *International Journal of Video & Image Processing and Network Security* 2009; 9: 291-297.
- [99] Szymanski JJ, Jamison JT, DeGracia DJ. Texture analysis of poly-adenylated mRNA staining following global brain ischemia and reperfusion. *Computer Methods and Programs in Biomedicine* 2012; 105: 81-94.
- [100] Filippi M, Horsfield MA, Campi A, Mammi S, Pereira C, Comi G. Resolution-dependent estimates of lesion volumes in magnetic resonance imaging studies of the brain in multiple sclerosis. *Annals of Neurology* 1995; 38: 749-754.
- [101] Skoch A, Jiráček D, Vyhnanovská P, Dezortová M, Fendrych P, Rolencová E, Hájek M. Classification of calf muscle MR images by texture analysis. *Magnetic Resonance Materials in Physics, Biology and Medicine* 2004; 16: 259-267.
- [102] Zhou F, Shiroishi M, Gong H, Zee CS. Multiple sclerosis: hyperintense lesions in the brain on T1-weighted MR images assessed by diffusion tensor imaging. *Journal of Magnetic Resonance Imaging* 2010; 31: 789-795.
- [103] Calabrese M, Rinaldi F, Seppi D, Favaretto A, Squarcina L, Mattisi I, Perini P, Bertoldo A, Gallo P. Cortical diffusion-tensor imaging abnormalities in multiple sclerosis: a 3-year longitudinal study. *Radiology* 2011; 261: 891-898.
- [104] Riccitelli G, Rocca MA, Pagani E, Rodegher ME, Rossi P, Falini A, Comi G, Filippi M. Cognitive impairment in multiple sclerosis is associated to different patterns of gray matter atrophy according to clinical phenotype. *Human Brain Mapping* 2011; 32: 1535-1543.
- [105] Fabiano AJ, Sharma J, Weinstock-Guttman B, Munschauer FE 3rd, Benedict RH, Zivadinov R, Bakshi R. Thalamic involvement in multiple sclerosis: a diffusion-weighted magnetic resonance imaging study. *Journal of Neuroimaging*, 2003; 13: 307-314.
- [106] Madabhushi A, Udupa JK. Interplay between intensity standardization and inhomogeneity correction in MR image processing. *IEEE Transactions on Medical Imaging* 2005; 24: 561–576.
- [107] Vovk U, Pernus F, Likar B. A review of methods for correction of intensity inhomogeneity in MRI. *IEEE Transactions on Medical Imaging* 2007; 26: 405-421.

- [108] Zhuge Y, Udupa JK, Liu J, Saha PK. Image background inhomogeneity correction in MRI via intensity standardization. *Computerized Medical Imaging and Graphics* 2009; 33: 7-16.
- [109] Woods BJ, Clymer BD, Kurc T, Heverhagen JT, Stevens R, Orsdemir A, Bulan O, Knopp MV. Malignant-lesion segmentation using 4D co-occurrence texture analysis applied to contrast-enhanced magnetic resonance breast image data. *Journal of Magnetic Resonance Imaging* 2007; 25: 495-501.
- [110] Mugler JP 3rd. Potential degradation in image quality due to selective averaging of phase-encoding lines in Fourier transform MRI. *Magnetic Resonance in Medicine* 1991; 19: 170-174.

ALKUPERÄISJULKAISUT

RESEARCH

Open Access

Effect of slice thickness on brain magnetic resonance image texture analysis

Sami J Savio^{1,2*}, Lara CV Harrison^{2,3}, Tiina Luukkaala^{4,5}, Tomi Heinonen^{2,6}, Prasun Dastidar^{1,2,3}, Seppo Soimakallio^{1,3}, Hannu J Eskola^{1,2}

* Correspondence: sami.savio@tut.fi
¹Medical Imaging Centre, Tampere
University Hospital, Biokatu 8,
Tampere, FI-33521, Finland

Abstract

Background: The accuracy of texture analysis in clinical evaluation of magnetic resonance images depends considerably on imaging arrangements and various image quality parameters. In this paper, we study the effect of slice thickness on brain tissue texture analysis using a statistical approach and classification of T1-weighted images of clinically confirmed multiple sclerosis patients.

Methods: We averaged the intensities of three consecutive 1-mm slices to simulate 3-mm slices. Two hundred sixty-four texture parameters were calculated for both the original and the averaged slices. Wilcoxon's signed ranks test was used to find differences between the regions of interest representing white matter and multiple sclerosis plaques. Linear and nonlinear discriminant analyses were applied with several separate training and test sets to determine the actual classification accuracy.

Results: Only moderate differences in distributions of the texture parameter value for 1-mm and simulated 3-mm-thick slices were found. Our study also showed that white matter areas are well separable from multiple sclerosis plaques even if the slice thickness differs between training and test sets.

Conclusions: Three-millimeter-thick magnetic resonance image slices acquired with a 1.5 T clinical magnetic resonance scanner seem to be sufficient for texture analysis of multiple sclerosis plaques and white matter tissue.

Background

Texture analysis (TA) is based on the examination of spatial patterns in image intensity. Many widely used texture analysis techniques exist in several fields of science, engineering and medical sciences. They have been successfully applied to several clinical applications, including multiple sclerosis (MS), brain injury and diseases that are otherwise difficult to identify at an early stage [1-3]. In neuroradiological imaging for clinical purposes, MS is the most common autoimmune disease of the central nervous system. It has a complex pathophysiology including inflammation, demyelination, axonal degeneration and neuronal loss. Diagnostic evaluation of MS is widely based on conventional magnetic resonance imaging (MRI) and the McDonald clinical diagnostic criteria [4,5]. The guidelines include evaluation of MS disease attacks, cerebrospinal fluid analysis and MRI. The MRI criteria includes three of the following 1) at least one active lesion seen on gadolinium(Gd)-enhanced T1 images, or if there is no Gd-enhancing lesion at least nine T2 hyperintense lesions; 2) At least one infratentorial lesion;

3) At least one juxtacortical lesion; 4) At least three periventricular lesions. Spinal cord lesions can be considered as equivalent brain lesions. In McDonald criteria the lesion size has no relevance on the inclusion criteria for MS. Only the number of lesions is important.

As interest in texture analysis has grown due to its wide range of useful applications, a thorough understanding of the impact of different physical factors on it remains incomplete. For instance, sequence selection, magnetic field strength, voxel size, image resolution, patient movement, sequence-based noise and slice thickness are factors that complicate the interpretation of the acquired data. Some of these factors will be shortly discussed below. However, in our current study we concentrate on the effects of slice thickness on texture analysis and especially in brain MR images.

Slice thickness

Herlidou-Meme et al. [6] have previously performed a multicenter texture analysis study in which three 1.5 T MRI units were used to acquire T1- and T2-weighted images with 2-mm, 4-mm and 6-mm slices for T1 and 2.5-mm, 5-mm and 7.5-mm slices for T2, respectively. Three classes of foam and gel test objects, ranging in size from 0.72-3.70 mm, were studied. According to their findings, the classification error was higher for small-scale objects, ranging from 10% for coarse foam to 40% for gel. They did not find the selection of slice thickness to be significant for classification accuracy, as the number of well-classified regions of interest (ROIs) seemed to be almost independent of slice thickness. However, their study included only a small set of objects.

Materka et al. [7] studied 1.5 T MR phantom images of foam-filled tubes with different fields of view (100 mm * 100 mm and 200 mm * 200 mm) and a constant number of image pixels (256 * 256) to find the effect of Gaussian noise and slice thickness (2 mm and 4 mm) on the Fisher coefficient F . Four texture classes were used with five samples in each. Only a few texture parameters with meaningful Fisher coefficient values were discovered. The number of useful parameters depended significantly on image normalization, as some texture parameters showed high correlation with the mean and variance. The Fisher coefficient decreased corresponding to the noise and slice thickness.

Guggenbuhl et al. [8] investigated the effect of slice thickness on texture parameters on computed tomography (CT) of calf bone images. They found that the selection of slice thickness used in the imaging had a significant influence on at least some run-length and co-occurrence parameters in the determination of bone microarchitecture. However, they did not perform any classification of textures in their study, and it remains unclear whether the classification accuracy of CT images would have been changed due to increased slice thicknesses.

Spatial resolution

Jirak et al. [9] studied polystyrene spheres of varying diameters, ranging from 0.8 to 2.0 mm, as well as an agar solution, to investigate texture analysis procedures on MR images. They found feature extraction techniques, such as the Fisher coefficient and the probability of error with an average correlation coefficient (POE+ACC), to be very sensitive to deviations in inter-pixel relationships, noise and inhomogeneity. High

resolution was often essential for the texture analysis of the analyzed materials. However, linear discriminant analysis (LDA) classification accuracy of objects with different spheres was over 90%, even in low-resolution images, thus suggesting that low-resolution images could be used for texture analysis. They also found that when high resolution was used the phantoms with small spheres and thus a more exact texture were easier to classify than those with larger ones, but with low resolution this was not true.

Mayerhoefer et al. [10] investigated the sensitivity of texture features to the variations in the number of acquisitions, repetition time (TR), time to echo (TE) and sampling bandwidth (SBW) in MR imaging using polystyrene spheres and agar gel phantoms. They found that texture features are increasingly sensitive to acquisition parameter variations as spatial resolution increases, but with a sufficient resolution, these variations did not have a considerable effect on the classification results in their study. As only relatively small test sets of spherical elements were used in these studies, it remains unclear whether the results can be fully generalized in the analyses of real acquisitions.

Signal-to-noise ratio

Schad and Lundervold [11] showed that the best discriminating features of different textures depend on the signal-to-noise ratio (SNR), spatial resolution and voxel size, comparable to the natural properties of the imaged regions. A few texture parameters provide relatively constant results with all studied SNR levels and voxel sizes, but there is usually a connection between voxel size and the texture parameter values. For example, some Angular Second Moment values seem to increase as a function of increasing voxel size. The exact relationship varies between parameters and studied tissues. The stabilization of texture parameter values as a function of increasing SNR was observed in grey matter, but in white matter, the stabilization was less systematic.

ROI selection

Harrison et al. [12] studied the sensitivity of texture analysis acquisition in MR images containing MS lesions at two different anatomical levels using Wilcoxon's signed ranks test. Several different tissues—white matter, normal appearing white matter, normal appearing grey matter, cerebrospinal fluid and MS plaques—were chosen for classification, and constant as well as manually drawn ROIs were used. According to the results, 96-100% of white matter (WM) or normal appearing white matter (NAWM) areas were classified correctly against MS plaques. No significant differences in the classification results for imaging sequences or anatomical levels were found in this study. The conclusion from this single center study was that texture analysis applied to MRI is a robust method when a fixed imaging sequence is used. However, they noted that the selection between a manually drawn and a rectangular ROI on an MS plaque depends on the specific application requirements and may have an impact on the accuracy of texture analysis.

The purpose of our study

In this paper we provide a study of the effect of MRI slice thickness on detecting MS lesions by means of texture analysis of brain tissues and structures. There are usually numerous small lesions in the brain parenchyma of an MS patient, which may not be

seen if the slice thickness is not small enough [13,14]. This means that the question of acquiring suitably thin slices and their effect on texture analysis plays an important role in the final diagnosis of the patient, especially as in earlier studies [7,8] it has been shown that many texture parameter values are dependable on the slice thickness.

If there are small lesions that are thin in axial direction, they may be mixed with the surrounding normal tissue present in thick slices. This underlines the importance of the slice thickness selection in the MR image acquisition. While many factors, such as the signal-to-noise ratio (SNR) and movement artifacts can influence the interpretation, we tried to make the analysis easier by averaging the intensities of three consecutive 1-mm slices to create a new slice, thus essentially expressing a 3-mm slice of the original subject.

The main aim of our study was to analyze the effect of slice thickness in the MRI examination for the detection of MS lesions using texture analysis. As far as we know, this effect has not been studied earlier by using real clinical material.

Methods

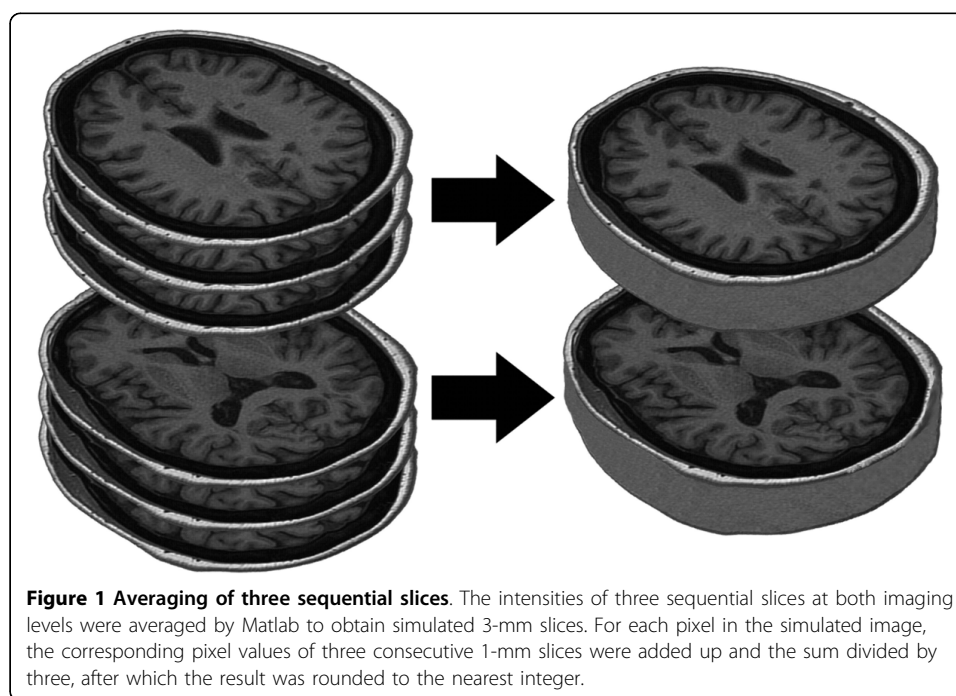
We investigated the effect of slice thickness on the results from texture analysis of brain MR images acquired using a 1.5 T MRI device (Siemens, Avanto Syngo MR B15, Erlangen, Germany). We studied MR images of 23 clinically diagnosed MS patients (13 females and 10 males; mean age $42 \pm \text{SD } 11$; age range from 18 to 60 years). The participants' neuroradiological status based on MRI examinations fulfilled the McDonalds revised diagnostic criteria for MS [4,5]. None of the patients underwent steroid therapy 3-6 before MRI examinations.

The image analysis was focused at two anatomical levels, the corona radiata and the centrum semiovale, along with MS lesions, as well as basal ganglia. The present study forms a part of an ongoing study where 100 patients with CIS, RRMS, PPMS and SPMS are being longitudinally studied with yearly clinical and MRI control examinations.

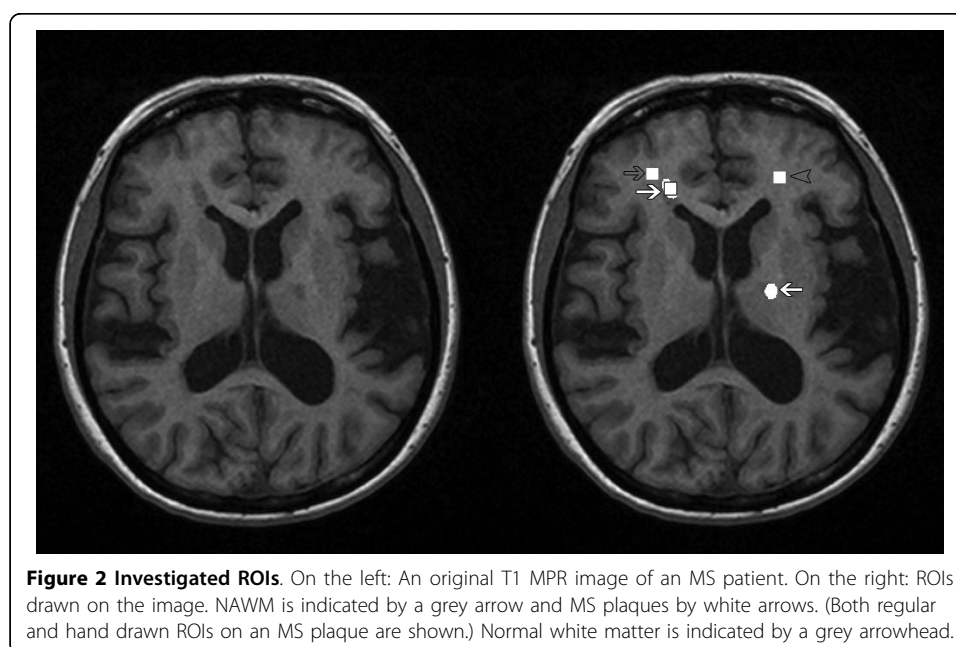
Imaging was performed by the hospital's clinical imaging protocol for MS follow-up. This included standardized axial T1-weighted 3D magnetization prepared gradient echo sequence (MPR) (denoted by "T1") and the previous T1-weighted MPR sequence with the intra-venous contrast agent Gadoterate meglumine (Gd-DOTA) (Dotarem® 10 ml) (denoted by "T1C"). The original slice thickness was 0.9 mm, which was interpreted as 1 mm for convenience. Other imaging parameters were field of view 230 mm*230 mm, TR 1160 ms, TE 4.24 ms, time for inversion (TI) 600 ms, echo train length 1, flip angle 15° and matrix size 512*512 pixels. To minimize imaging artifacts, a pre-scan normalization filter was used for the intensity inhomogeneity correction in the images. According to an experienced radiologist, the image sets were qualified as uniform, with good quality and free from motion artifacts.

The simulated 3-mm slices were created with Matlab (v. 7.7.0.471) [15] by intensity averaging of three sequential 1-mm slices, with the middle slice belonging to the studied set of 1-mm slices, see Figure 1. Each pixel intensity value in the simulated image was an average of the corresponding pixel intensity values of the above mentioned three consecutive 1 mm slices, rounded to the nearest integer.

Four ROIs were drawn in each original and simulated image: normal white matter (denoted by "WM"), normal appearing white matter adjacent to an MS plaque ("NAWM"), a manually drawn irregular ROI (size between 39 and 2477 pixels, medium



size of 354 pixels) over an MS plaque ("MSi") and a constant rectangular ROI (size of 10*10 pixels) over an MS plaque ("MSr"). Rectangular ROI size was defined to fit on the average sized MS plaque. Image grey levels were normalized to limit image intensities to between $[\mu-3\sigma, \mu+3\sigma]$ for each ROI, where μ is the mean grey level, and σ is the standard deviation. This normalization scheme was reported by Collewet et al. [16] to result in the best classification results in the case of MR imaging of soft cheeses and having no relationship between classification errors and MR acquisition protocols. An example of ROIs drawn is shown in Figure 2.



For each ROI, we calculated 264 texture parameter values by MaZda (v. 3.20) [17], a toolkit created by the European COST project [1]. These included 11 histogram parameters representing the gray level frequencies, 4 absolute gradient parameters showing neighboring pixel variation, 220 co-occurrence matrix parameters measuring probabilities of pixel pair occurrences, 20 run-length parameters to define frequencies of defined length and gray level, 4 wavelet and 5 autoregressive model-based parameters.

We made no assumptions for the data distribution and used Wilcoxon's signed ranks test to calculate a statistical measure for all texture parameters to find differences between ROIs. The resulting p-values that were smaller than 0.05 were considered as statistically significant, and the proportion of significant p-values was used as a separability measure.

We also tested the actual classification separability of each tissue pair by linear and nonlinear discriminant classifiers with Matlab [15]. If the training and test sets are not separate or they are small, there is a possibility that the classification results are too optimistic, or even that they are purely a coincidence in some cases. To avoid these risks and to take into account the fact that high correlation would probably exist between several texture parameters, we limited the amount of selected parameters for classifiers to three for each classification task by the Fisher coefficient. The Fisher coefficient was calculated in Matlab for all 264 texture parameters. The three parameters giving highest Fisher coefficient values for each classification task were included in the further classification procedures with linear and nonlinear discriminant classifiers. The Fisher coefficient F is defined as follows [7]:

$$F = \frac{D}{V},$$

where

D = between-classes variance and

V = within-classes variance.

When running linear and nonlinear discriminant classifiers, one hundred classifications with separate training and test sets were run for each comparison, and the median result was used as the measure for the classification accuracy. When the slice thicknesses for the training and test sets were different, we also made sure that the data from the same patient at the same physiological location did not belong to the training and test sets at the same time.

Results

Tissue separability based on Wilcoxon's signed ranks test

We calculated the texture parameter values for each combination of image sequences (T1 and T1C) and slice thicknesses (1 mm and simulated 3 mm). We also applied Wilcoxon's signed ranks test to these values. In the following analyses, each given separability measure is the percentage of statistically significant texture parameters when two ROIs are compared with a fixed slice thickness. In the following tables, "1 mm" refers to the original slices and "3 mm" to the simulated 3-mm slices. Table 1 shows the statistical separability between tissues based on Wilcoxon's signed ranks test.

Table 1 The percentage of statistically significant parameters

ROIs	1 mm		3 mm	
	T1	T1C	T1	T1C
WM vs. NAWM	12	1	11	11
WM vs. MSi	80	84	81	80
WM vs. MSr	76	74	78	84
NAWM vs. MSi	80	84	79	82
NAWM vs. MSr	75	78	77	80
MSi vs. MSr	66	73	71	73

The percentage of statistically significant parameters ($p < 0.05$) with slice thicknesses of 1 mm and 3 mm in Wilcoxon's signed ranks test for T1 and T1C (with contrast agent) images in comparison with white matter (WM), normal appearing white matter (NAWM), MS plaques with irregular ROI (MSi) and MS plaques with regular ROI (MSr). Each comparison includes 264 texture parameters.

Statistically significant differences were observed between tissues in the brain, both in the analyses of different slice thicknesses and in that of MRI sequences. Only the difference between WM and NAWM did not show high statistical significance.

We also calculated the Fisher coefficient for each texture parameter to select parameters for classification. Texture parameters that were chosen most often for the analysis by the Fisher coefficient value with sequence T1 were sigma and correlation at a distance of five pixels and a direction of 135 degrees (S(5,-5)Correlat), whereas the most used parameters for sequence T1C were sigma and sum variance at a distance of three pixels and a direction of 135 degrees (S(3,-3)SumVarnc). However, the chosen texture parameters depended on the compared ROIs as well as the slice thickness in the training set.

Tissue classification results

The results of the linear discriminant analysis (LDA) between ROIs (excluding the MSi vs. MSr comparison) are shown in Table 2.

The classification results selected by linear discriminant analysis show that tissue classification can be performed reliably using both 1-mm and simulated 3-mm slices. This was applicable in the training set as well as in the test set.

The results of the nonlinear discriminant analysis (NDA) between ROIs (excluding MSi vs. MSr comparison) are shown in Table 3.

The performed nonlinear discriminant analyses show results similar to those presented in Table 2.

The effect of slice thickness on texture parameter distributions

For each region of interest, we compared the texture parameter values for the original 1-mm slice and the simulated 3-mm slice in T1 and T1C images. The comparison was made by indicating the percentage of statistically significant p-values of 264 texture parameters. The results are shown in Table 4.

Intra-tissue comparisons with different slice thicknesses show statistical differences in texture parameter values. The texture parameter values with white matter show less statistically significant changes compared to that with MS plaques.

Discussion

The aim of this study was to find direct applications for clinical analysis of TA in the field of MS. Our study included the evaluation of the effect of different slice

Table 2 The accuracy of LDA method for T1 and T1C

ROIs	1 mm -> 1 mm		1 mm -> 3 mm		3 mm -> 1 mm		3 mm -> 3 mm	
	T1	T1C	T1	T1C	T1	T1C	T1	T1C
WM vs. NAWM	60 [59, 61]	55 [53, 56]	52 [50, 54]	50 [48, 53]	57 [55, 58]	55 [53, 56]	60 [59, 61]	62 [60, 63]
WM vs. MSi	92 [91, 95]	98 [97, 98]	93 [93, 94]	93 [92, 93]	92 [91, 95]	89 [89, 90]	96 [95, 96]	95 [94, 95]
WM vs. MSr	93 [93, 94]	93 [92, 94]	88 [88, 89]	91 [91, 93]	90 [90, 91]	84 [83, 85]	89 [88, 89]	92 [92, 93]
NAWM vs. MSi	91 [90, 92]	98 [97, 98]	92 [92, 93]	98 [98, 98]	95 [94, 96]	98 [98, 100]	95 [94, 96]	100 [98, 100]
NAWM vs. MSr	91 [90, 92]	90 [89, 91]	88 [86, 88]	90 [89, 90]	90 [90, 91]	91 [90, 92]	88 [86, 88]	94 [93, 94]

Table 3 The accuracy of NDA method for T1 and T1C

ROIs	1 mm -> 1 mm		1 mm -> 3 mm		3 mm -> 1 mm		3 mm -> 3 mm	
	T1	T1C	T1	T1C	T1	T1C	T1	T1C
WM vs. NAWM	59 [57, 60]	50 [48, 52]	54 [52, 56]	48 [46, 50]	54 [54, 55]	50 [50, 53]	57 [55, 58]	60 [59, 63]
WM vs. MSi	95 [94, 96]	96 [94, 97]	94 [93, 95]	92 [91, 93]	92 [91, 93]	88 [88, 89]	96 [96, 98]	93 [92, 94]
WM vs. MSr	92 [92, 93]	92 [91, 93]	85 [85, 86]	92 [91, 92]	90 [89, 90]	81 [80, 82]	88 [88, 89]	92 [91, 92]
NAWM vs. MSi	92 [91, 92]	98 [96, 98]	91 [90, 92]	97 [96, 98]	91 [90, 93]	98 [97, 98]	95 [93, 96]	98 [98, 100]
NAWM vs. MSr	90 [89, 90]	88 [88, 89]	86 [85, 88]	90 [89, 91]	89 [88, 90]	90 [89, 91]	86 [85, 86]	92 [91, 92]

Table 4 Intra-tissue comparison between slice thicknesses

ROI	T1	T1C
WM	34	45
NAWM	44	42
MSi	57	67
MSr	51	59

Intra-tissue comparison between slice thicknesses. The percentage of statistically significant p-values ($p < 0.05$) when each region of interest (white matter (WM), normal appearing white matter (NAWM), MS plaques with irregular ROI (MSi) and MS plaques with regular ROI (MSr)) on a 1-mm slice is compared to the respective ROI on a 3-mm slice in T1 and T1C (with contrast agent) images. The comparison was made by indicating the percentage of statistically significant p-values of 264 texture parameters.

thicknesses in MRI examination for detecting MS lesions using texture analysis. Our results show that only minor differences exist between different slice thicknesses. This can be clearly seen when the LDA and NDA classification results are analyzed.

Two-dimensional vs. three-dimensional analysis

In this study two-dimensional analysis was performed even though three-dimensional texture analysis of brain images, applied by Mahmoud-Ghoneim et al. [18], was found to be a more exact method mathematically than its two-dimensional counterpart. However, the clinical evaluation remains mostly based on two-dimensional slices, and a fast two-dimensional texture analysis is easier to combine with this procedure. Furthermore, the slice thickness is normally much larger than the spatial resolution in each plane, and thus, non-cubical voxels are produced. For the clinician, it is also easier, faster and more reliable to draw ROIs two-dimensionally. In some cases only single slices are available, thus making three-dimensional analysis impossible.

Data acquisition

In a sense it would have been more practical to acquire original 3 mm slices and compare their texture to the 1 mm ones. Our choice was a limited one regarding this because the MRI examinations of the MS patients studied here were already performed earlier using 1 mm thick slices, making our examination a retrospective one.

We synthesized thicker slices by averaging consecutive original slices. Both slice types were then studied to reveal the effect of slice thickness variation on texture analysis. Clearly, not all tissue properties are preserved in the averaging process and the real partial volume effect might have been a bit different. However, a clear advantage of this simple method is that even though the slice thickness was changed, the other basic parameters were left unchanged. The differences can thus be interpreted as the effect of thicker slices.

Axial resolution

Brain tissue is not axially symmetrical, which is challenging for slice averaging. If several bones, muscles, or internal organs were to be studied, the differences between consecutive slices would probably be smaller. As Guggenbuhl et al. have noted [8], texture parameter values change when the slice thickness is altered. However, the influence of slice thickness on classification results seems to be relatively small in texture analysis of MS. The axial resolution of MR images decreases and SNR increases as the slices becomes thicker. These two factors have opposite effects on the accuracy of

texture analysis, and their exact impact is case dependent. In the following subsections, we discuss the separability of tissues when two slice thicknesses are used.

Statistical comparison of texture parameter values between ROIs

Based on our results, WM and NAWM seem to have the smallest mutual separability. This is true for both imaging sequences as well as for both slice thicknesses. A low value of separability was expected, as according to earlier studies [12,19], WM and NAWM are relatively hard to distinguish from one another with the clinically suitable 1.5 T magnetic field strength. Statistically significant p-values were found in only 11-12% of cases with T1 and in 1-11% with T1C (Table 1). Some of these values may even partly arise from the statistical significance level of 0.05 we have used.

When sequence T1 is considered, the ROIs that differ the most in statistical terms are WM compared to MSi (80-81% of p-values were statistically significant), NAWM to MSi (79-80%), WM to MSr (76-78%) and NAWM to MSr (75-77%). Even the ROI comparison of MSi to MSr (66-71%) provides a high proportion of texture parameters with statistically significant, low p-values. With sequence T1C, NAWM and MSi (82-84%) as well as WM and MSi (80-84%) are now the easiest ROIs to separate, followed by WM and MSr (74-84%) and NAWM and MSr (78-80%). In addition, the ROIs MSi and MSr are quite distinct from one another (73%) (Table 1).

When manually drawn ROIs are compared to the fixed size rectangular ROI on MS plaques it is important to note that manually drawn ROIs do not overlap with the neighboring tissues, whereas differences in lesion sizes between individuals lead to a variable amount of partial volume pixels in the standard-sized ROI boxes. Harrison et al. [12] expected this matter to have some impact on the accuracy of texture analysis. However, they state that standardized ROIs have several other advantages, including increased tolerance for slice selection and reproducibility.

We may also postulate that the average separability measure does not greatly depend on the selected imaging sequence or the slice thickness used, and the order of the separability values seems to be quite natural as well. White matter areas are very different compared to MS plaques. The separability of MS plaques from both WM and NAWM is high.

These findings are positive and suggest that many possible parameters are effective for the MS texture analysis and are independent of the slice thicknesses (1 mm or 3 mm) used in this study. In addition, in the study reported by Herlidou-Meme et al. [6], the classification results for foam were obtained almost independently of the slice thickness. It should be kept in mind, however, that in the mentioned study, the test set was relatively small.

When we compared WM and NAWM, the tissues were generally very difficult to separate; it also seems that the use of Gadoterate meglumine as a contrast agent does not have a significant influence on the texture classification accuracy.

Tissue classification

We used linear discriminant analysis to test the real effect of the separability deduced by Wilcoxon's signed ranks test. The classification results show that MS and WM areas are well separable in terms of LDA classification. With some exceptions, WM and MS plaques were classified with at least 90% accuracy, as evident in Table 2. The

classification with simulated 3-mm slices is almost as accurate as with 1-mm slices. The effect of sequence selection is rather small.

The results remain almost the same even if the discrimination decision was created based on the training set consisting of 1-mm slices and applied to simulated 3-mm slices in the test set, or vice versa. The order of classification accuracy by this method is also nearly the same as that suggested by Wilcoxon's signed ranks test; NAWM vs. MSi (91-100%), WM vs. MSi (89-98%), NAWM vs. MSr (88-94%) and WM vs. MSr (84-93%); therefore, the regular ROIs are harder to classify than the irregular ones.

The difference between 1-mm and simulated 3-mm slices in the linear classification of MS plaques and white matter thus seems rather small. One can also determine that the classification accuracy between NAWM and WM is poor, as could be expected from the separability results shown in Table 1.

We also tested NDA on the same ROIs and found that the classification results were nearly the same as in the case of LDA. From the results shown in Table 3, we see that, as in the case of LDA, the classification accuracy of NDA is also 90% or more in most cases, where WM or NAWM are classified against MS ROIs, and the order of the classification accuracy remains the same as that of LDA: NAWM vs. MSi (91-98%), WM vs. MSi (88-96%), NAWM vs. MSr (86-92%) and WM vs. MSr (81-92%). The classification accuracy between NAWM and WM is also poor with NDA.

Comparison of texture parameter values between slice thicknesses

The selection of slice thickness seems to have a greater effect on the texture analysis of MS plaques than in the white matter areas. The results in Table 4 show the average statistical differences between texture parameter values calculated for 1-mm and simulated 3-mm slices. For both sequences T1 and T1C, the amount of statistically significant p-values is close to 50%, averaging 46% for T1 and 54% for T1C, when calculated over all four ROIs studied.

This separability measure ranges from 34% (WM) to 57% (a manually drawn irregular ROI over an MS plaque) when sequence T1 is considered, and from 42% (NAWM) to 67% (a manually drawn irregular ROI over an MS plaque) when T1C is examined. Textural differences in WM areas between 1-mm and simulated 3-mm slices seem to be decreased compared to those of MS plaques. These differences do not seem to affect classification results, as evident from the actual classification results shown in Tables 2 and 3. In general, there seems to be some statistical difference between 1-mm and 3-mm texture parameter distributions. The same effect in bone CT imaging was reported earlier by Guggenbuhl et al. [8]

General findings

The interpretation of our findings is that the results from texture analysis vary slightly if the slice thickness is increased from 1-mm to 3-mm. As indicated by Schad and Lundervold [11], for example, textures also depend on spatial resolution and voxel size, and thus, it is important to investigate in which circumstances the texture can be identified. Changes in the texture due to the increased slice thickness are usually still small enough to enable adequate texture classification between MS plaques and white matter areas when suitable texture parameters are chosen. However, depending on the exact application, separate texture parameter sets for 1-mm and 3-mm slices may have to be

used to enable sufficiently accurate classification. Although our findings are aligned with earlier studies, further clarification is still needed to understand the dependence of texture classification accuracy on slice thickness.

Conclusions

In this study we found that replacing 1-mm slices with 3-mm-thick slices does not remove textures which would prevent a clinician from detecting the presence of lesions visible on 1-mm-thick slices. It is clear that very thin slices would produce insufficient SNR for reliable texture analysis, whereas very thick slices lack the texture details. See-ingly, the range from 1 mm to 3 mm does not correspond to either of these cases. To apply texture analysis in the quantification of clinical images, further studies are needed to show the effect of slice thickness as well as other parameters, such as move-ment artifacts and noise, on its performance.

Acknowledgements

The South Savo Regional Fund of the Finnish Cultural Foundation and the Tampere City Science Fund supported this research.

Author details

¹Medical Imaging Centre, Tampere University Hospital, Biokatu 8, Tampere, FI-33521, Finland. ²Department of Biomedical Engineering, Tampere University of Technology, Korkeakoulunkatu 3, Tampere, FI-33720, Finland. ³Tampere University Medical School, University of Tampere, Medisiininkatu 3, Tampere, FI-33014, Finland. ⁴Science Center, Pirkanmaa Hospital District, Biokatu 12, Tampere, FI-33521, Finland. ⁵Tampere School of Public Health, University of Tampere, Medisiininkatu 3, Tampere, FI-33014, Finland. ⁶Department of System Engineering and Architecture, Nokia Corporation, Visiokatu 1, Tampere, FI-33720, Finland.

Authors' contributions

SJS participated in the design of the study, carried out the statistical image analysis and drafted the manuscript. LCVH designed the image analysis and participated in the statistical analysis and interpretation of the results. TL participated in the design of the study and helped to perform the statistical analysis. TH participated in the study design and coordination. PD participated in design and coordination as well as result interpretation. SS participated in the design and coordination of the study. HJE participated in the design, coordination and result interpretation. All authors participated in drafting the manuscript and read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 4 May 2010 Accepted: 18 October 2010 Published: 18 October 2010

References

1. Hajek M, Dezortova M, Materka A, Lerski R, (Ed): **Texture Analysis for Magnetic Resonance Imaging**. Prague, Med4 Publishing 2006.
2. Castellano G, Bonilha L, Li LM, Cendes F: **Texture analysis of medical images**. *Clin Radiol* 2004, **59**:1061-1069.
3. Bonilha L, Kobayashi E, Castellano G, Coelho G, Tinois E, Cendes F, Li LM: **Texture analysis of hippocampal sclerosis**. *Epilepsia* 2003, **44**:1546-1550.
4. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, McFarland HF, Paty DW, Polman CH, Reingold SC, Sandberg-Wollheim M, Sibley W, Thompson A, Van Den Noort S, Weinshenker BY, Wolinsky JS: **Recommended diagnostic criteria for multiple sclerosis: Guidelines from the international panel on the diagnosis of multiple sclerosis**. *Ann Neurol* 2001, **50**:121-127.
5. Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, Lublin FD, Metz LM, McFarland HF, O'Connor PW, Sandberg-Wollheim M, Thompson AJ, Weinshenker BG, Wolinsky JS: **Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald criteria"**. *Ann Neurol* 2005, **58**:840-846.
6. Herlidou-Meme S, Constans JM, Carsin B, Olivier D, Eliat PA, Nadal-Desparats L, Gondry C, Le Rumeur E, Idy-Peretti I, de Certaines JD: **MRI texture analysis on texture test objects, normal brain and intracranial tumors**. *Magn Reson Imaging* 2003, **21**:989-993.
7. Materka A, Strzelecki M, Lerski R, Schad L: **Feature evaluation of texture test objects for magnetic resonance imaging**. In *Texture Analysis in Machine Vision, Series in Machine Perception and Artificial Intelligence*. Edited by: Pietikainen MK. Singapore, World Scientific Publishing Co. Pte. Ltd; 2000:**40**:197-206.
8. Guggenbuhl P, Chappard D, Garreau M, Bansard JY, Chales G, Rolland Y: **Reproducibility of CT based bone texture parameters of cancellous calf bone samples: Influence of slice thickness**. *Eur J Radiol* 2008, **67**:514-520.
9. Jirak D, Dezortova M, Hajek M: **Phantoms for texture analysis of MR images**. In *Texture Analysis for Magnetic Resonance Imaging*. Edited by: Hajek M, Dezortova M, Materka A, Lerski R. Prague, Med4 Publishing; 2006:115-127.

10. Mayerhoefer ME, Szomolanyi P, Jirak D, Materka A, Trattnig S: **Effects of MRI acquisition parameter variations and protocol heterogeneity on the results of texture analysis and pattern discrimination: An application-oriented study.** *Med Phys* 2009, **36**:1236-1243.
11. Schad LR, Lundervold A: **Influence of resolution and signal to noise ratio on MR image texture.** In *Texture Analysis for Magnetic Resonance Imaging*. Edited by: Hajek M, Dezortova M, Materka A, Lerski R. Prague, Med4 Publishing; 2006:129-149.
12. Harrison LCV, Raunio M, Holli K, Luukkaala T, Savio S, Elovaara I, Soimakallio S, Eskola H, Dastidar P: **MRI texture analysis in multiple sclerosis: Toward clinical analysis protocol.** *Acad Radiol* 2010, **17**:696-707.
13. Barbosa S, Blumhardt LD, Roberts N, Lock T, Edwards RHT: **Magnetic resonance relaxation time mapping in multiple sclerosis: normal appearing white matter and the 'invisible' lesion load.** *Magn Reson Imaging* 1994, **12**:33-42.
14. Filippi M, Horsfield MA, Campi A, Mammi S, Pereira C, Comi G: **Resolution-dependent estimates of lesion volumes in magnetic resonance imaging studies of the brain in multiple sclerosis.** *Ann Neurol* 1995, **38**:749-754.
15. **MATLAB - The Language Of Technical Computing.** [<http://www.mathworks.com/products/matlab/>].
16. Collewet G, Strzelecki M, Mariette F: **Influence of MRI acquisition protocols and image intensity normalization methods on texture classification.** *Magn Reson Imaging* 2004, **22**:81-91.
17. **MaZda Web Site.** [<http://www.elel.p.lodz.pl/programy/mazda/>].
18. Mahmoud-Ghoneim D, Toussaint G, Constans JM, de Certaines JD: **Three dimensional texture analysis in MRI: A preliminary evaluation in gliomas.** *Magn Reson Imaging* 2003, **21**:983-987.
19. Zhang J, Tong L, Wang L, Li N: **Texture analysis of multiple sclerosis: A comparative study.** *Magn Reson Imaging* 2008, **26**:1160-1166.

doi:10.1186/1475-925X-9-60

Cite this article as: Savio et al.: Effect of slice thickness on brain magnetic resonance image texture analysis. *BioMedical Engineering OnLine* 2010 **9**:60.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Estimation of sufficient signal to noise ratio for texture analysis of magnetic resonance images

Sami Savio^{*a,b}, Lara Harrison^{a,b}, Pertti Ryymin^a, Prasun Dastidar^a, Seppo Soimakallio^a, Hannu Eskola^{a,b}

^aMedical Imaging Center, Tampere University Hospital, Teiskontie 35, Tampere, FI-33521, Finland

^bDept. of Biomedical Engineering, Tampere University of Technology, Korkeakoulunkatu 3, Tampere, FI-33720, Finland

ABSTRACT

In this study, we have studied the effect of background noise on the texture analysis of muscle, bone marrow and fat tissues in 1.5 T magnetic resonance (MR) images using different statistical methods. Variable levels of noise were first added on 3-mm thick T2 weighted image slices of voluntary subjects to simulate several signal-to-noise ratio (SNR) levels. For each original and simulated image, the values for 264 texture parameters were calculated using MaZda, a texture analysis toolkit. We also determined Fisher coefficients based on the texture parameter values in order to enable high discrimination between different tissues. Linear discriminant analysis (LDA) and two different nearest neighbour (NN) methods were then applied for the texture parameters with the highest Fisher coefficient values. Several training and test sets were used to approximate the variation in the classification results. All the above-mentioned methods had the same classification accuracy, which in turn depended on the image SNR. We conclude that these tissues can be detected by texture analysis methods with a sufficient accuracy (90%) especially if SNR is at least 30-40 dB, even though the separation of different muscles remains a very challenging task.

Keywords: Texture analysis, signal-to-noise ratio, background noise, Fisher coefficient

1. INTRODUCTION

Texture analysis is a promising tool for image analysis, ranging from detecting abnormal tissues to therapy effect follow-up. Several applications of texture analysis can be found in the literature [1]. To do such processing, the quality of the original images must be sufficient. Movement artefacts, slice thickness and noise level have critical limits that should not be exceeded when applying texture analysis. Our research group has earlier studied the influence of slice thickness on texture analysis [2]. In this paper, we concentrate on the effect of noise levels on texture analysis.

Signal-to-noise ratio is a very important measure of image quality. A proper level of SNR is critical for several applications of signal processing. For example, in the field of diffusion tensor imaging (DTI) it has been shown that eigenvectors are systematically shifted and certain eigenvalues overestimated when the SNR is low, as it often is in in vivo measurements [3]. Also Farrell et al. [4] have investigated the effect of SNR on the results of DTI. They note that the values of fractional anisotropy measurements were biased due to decreased SNR, but the same effect was not perceived with mean diffusivity values. Additionally, fractional anisotropy quantification in brain grey matter was more difficult than in white matter when noise level was high. Erdogmus et al. [5] have emphasized the need for careful judgment as to whether SNR can be used as an objective measure of image quality. According to them, in certain nonlinear operations true SNR and measured SNR behave differently, either of them decreasing and the other increasing.

To study the effect of Gaussian noise and slice thickness on Fisher coefficient, Materka et al. [6] have investigated 1.5 T MR phantom images of foam filled tubes. They examined four texture classes, each consisting of five samples. The authors defined 6 as a threshold for an adequate value of Fisher coefficient. In their study, the Fisher coefficient of only a few texture parameters reached this threshold value. When they analyzed the Fisher coefficients of autoregressive model parameters further, they found them decreasing when noise was increased. According to the authors, the number of texture parameters that could be used for analysis depended significantly on the image normalization method, because some parameters had a high correlation with mean and variance.

*sami.savio@pshp.fi; phone +358 452 608 308; www.pshp.fi

Jirak et al. have studied texture analysis procedures on MR images [7]. Phantoms consisting of polystyrene spheres of various diameters and agar gel were imaged by T1 and T2 weighted MR. Small deviations in inter-pixel relationships due to noise, non-homogeneity and other factors led to a variation in values provided by the Fisher coefficient and the probability of error with average correlation coefficient (POE+ACC) method. If the resolution was high enough, classification accuracy of spherical objects was 90-100%, whereas low resolution often led to poor accuracy.

In an earlier phantom examination [8], the highest Fisher coefficients, depending on the SNR level, were much higher for polystyrene spheres than for foam and gel. The " $\pm 3\sigma$ normalization" method was used before the texture parameter calculation. Two texture parameters according to Fisher criteria were selected for the classification of polystyrene objects as well as two separate parameters for synthetic foam classification. For high SNR test cases, the nearest neighbour (1-NN) classification was perfect. As SNR levels decreased, the classification error increased. With the lowest SNR level, classification error was close to 15% for polystyrene and 37% for foam. According to this investigation, insufficient spatial resolution may be the reason for the difference.

In a study by Schad and Lundervold, SNR, spatial resolution, and voxel size had an effect on the most discriminating features of different textures in MR images of the brain [9]. They regularly found a connection between texture parameter values and chosen voxel size, although some parameters did not change much between the studied SNR levels and varying resolution. They also tested the classification accuracy based on four separate SNR levels. The images with the highest SNR level showed a high classification accuracy (more than 97% in the test set), when the lowest SNR test set gave only 62% accuracy. The most discriminating texture parameters changed when SNR decreased.

Skoch et al. [10] have studied T1 and T2 weighted MR images of healthy and diseased calf muscles. They found that 80% of the classification results agreed with the findings made by radiologists. However, the subjects with the least severe diseases were not always separated from the control group. The classification of muscles using texture analysis was not an easy task, even though there was a wide range of available texture parameters.

In this study, we have analyzed the maximum amount of noise in MR images that texture analysis can tolerate and yet obtain sufficiently accurate results when classifying real human tissues. Some of the tissues used are very different from the others; some in turn are very similar. Our data consists of thigh images, to which noise was added in order to create several SNR levels, grouped into distinct classes. The goal of our study was to analyze the sensitivity of texture analysis to noise.

2. MATERIALS AND METHODS

In the study, we used T2-weighted MR thigh images from three subjects. MRI was performed using a 1.5 T MRI device (Siemens, Avanto, version Syngo MR B15, Erlangen, Germany). Image intensity inhomogeneities were minimized with the use of a prescan normalization filter. Axial T2*-weighted gradient echo volumetric sequence (TR 36 ms, TE 15 ms, slice thickness 3.00 mm, pixel size 0.81 mm x 0.81 mm, flip angle 12°) was used.

Six separate imaging levels of each healthy volunteer were taken into consideration, in order to achieve the different SNR levels acquired from a real imaging situation. As the quality of the original images was high in the middle of the series but deteriorated at both ends, all images used in this study were selected from the middle of the series. Separate ROIs were drawn in bone marrow, two muscles: musculus rectus femoris and musculus vastus lateralis and subcutaneous fat. These tissue types were chosen to give a good combination of the different tissues that are subject to texture analysis in clinical applications. The above-mentioned tissues also represent homogenous as well as heterogeneous textures. The two selected muscles represent the same histological type of tissue, skeletal muscles. These muscles have a clear direction, and the muscles are parallel in the thigh area. Axial MR images represent almost exactly the cross sections of these muscles. The muscle texture is visually homogenous and smooth. Visually, these muscles can not be discriminated based on their textural appearance, this probably also makes quantitative classification a challenging task. The structure and texture of subcutaneous fat is heterogeneous and granular. The bone marrow is more fine-grained. Visually, different tissue types are detectable from images. An example of the ROIs drawn on these tissues can be seen in Figure 1.

To each of the original 18 images, seven separate levels of artificial Gaussian noise have been added in k-space. This leaves the intensity distribution of the resulted images Rician, as explained by Dietrich et al. [11] Complex Gaussian

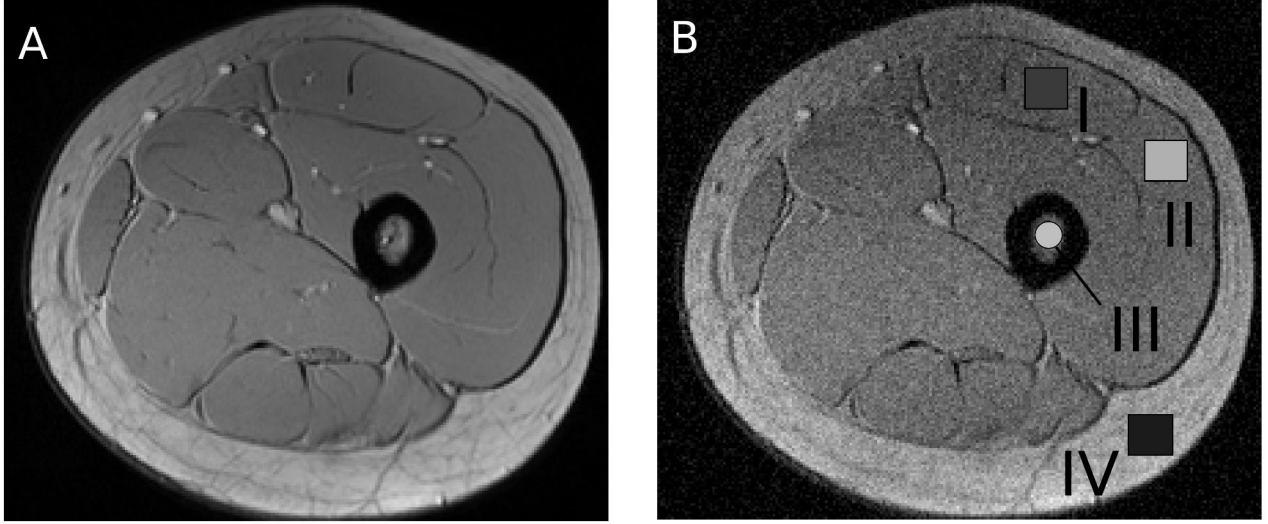


Figure 1. A) Original image of the thigh tissue. B) The following tissues have been selected in a noisy image: m. rectus femoris (I), m. vastus lateralis (II), bone marrow (III) and fat (IV).

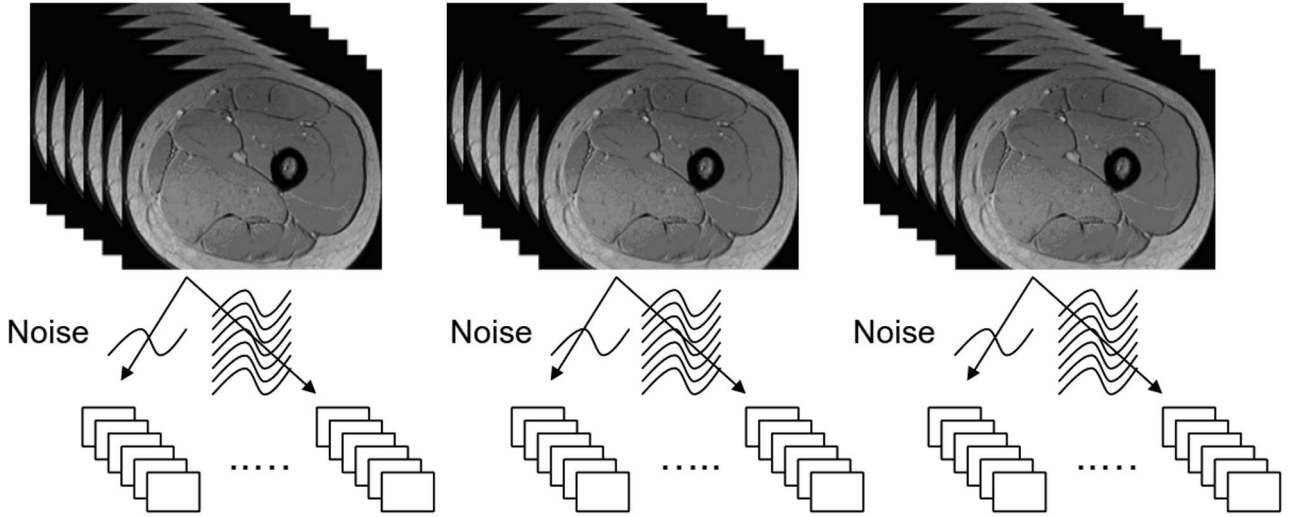


Figure 2. 18 MR images of voluntary subjects were modified to simulate several SNR levels. Seven different levels of noise were added in each original MR image, resulting in 18 original and 126 simulated images.

noise with variance σ^2 added to the original signal S leads to the following Rician probability distribution of the signal intensity S^* :

$$P(S^*; S, \sigma^2) = \frac{S^*}{\sigma^2} e^{-\frac{S^{*2} + S^2}{2\sigma^2}} I_0\left(\frac{S^* S}{\sigma^2}\right) \quad (1)$$

where I_0 is the modified Bessel function of zeroth order. This addition simulates real noise appearing in MR images.

For each ROI, eight different SNR levels were obtained for every image, as shown in Figure 2. The resulting 144 images have been divided into six classes with the same number of images, 24 in each, depending on the SNR. Table 1 shows

Table 1: Lower and upper SNR limits (dB) for tissue specific SNR classes (1-6) for m. rectus femoris, m. vastus lateralis, fat and bone marrow.

SNR class	Musculus rectus femoris	Musculus vastus lateralis	Fat	Bone marrow
1	1-13	3-15	0-11	0-6
2	14-18	16-20	12-18	7-14
3	19-24	21-27	19-24	15-20
4	25-30	28-32	25-32	21-29
5	31-36	33-37	33-39	30-36
6	37-41	38-42	40-47	37-43

the minimum and maximum SNRs for each ROI and SNR class used in this study. $\pm 3\sigma$ normalization was applied on the images, after which Fisher coefficients for 264 texture parameters were calculated for each SNR class, including 220 co-occurrence matrix, 20 run-length matrix, 5 autoregressive model based, 4 absolute gradient and 4 wavelet based texture parameters originated by the texture analysis toolkit MaZda [12].

As the intensity range of separate tissue types are seemingly very different from each other but also vary between images, we did not use histogram based parameters in this study. Next, only three texture parameters with the highest Fisher coefficients have been chosen in each classification, thus decreasing the possibility of overtraining.

Finally, we measured the classification accuracy for each SNR class with linear discriminant analysis (LDA) and k-nearest neighbours (k-NN) classifiers. These classifiers have been trained by training sets and tested by separate test sets. Each test set and each training set was created by taking one half of the ROIs (12 of each) belonging to the SNR class in question. Different elements were always used in the corresponding training and test sets. Training and testing was performed 1000 times, and the median value of all these classification accuracies was selected. The repetition of classification also enabled estimation of 95% confidence intervals.

3. RESULTS

Classification accuracies of LDA and k-NN classifiers ($k = 1$ and 3) between m. rectus femoris, m. vastus lateralis, fat and bone marrow are shown in Figure 3, Figure 4 and Figure 5. In each test case, we have used either SNR class 5 or SNR class 6 as a training set. The SNR classes used for testing have varied from 1 to 6. We have not used any histogram based parameters to make the analysis more reliable. Because confidence intervals are usually small, only median confidence intervals greater than or equal to 3% units have been shown.

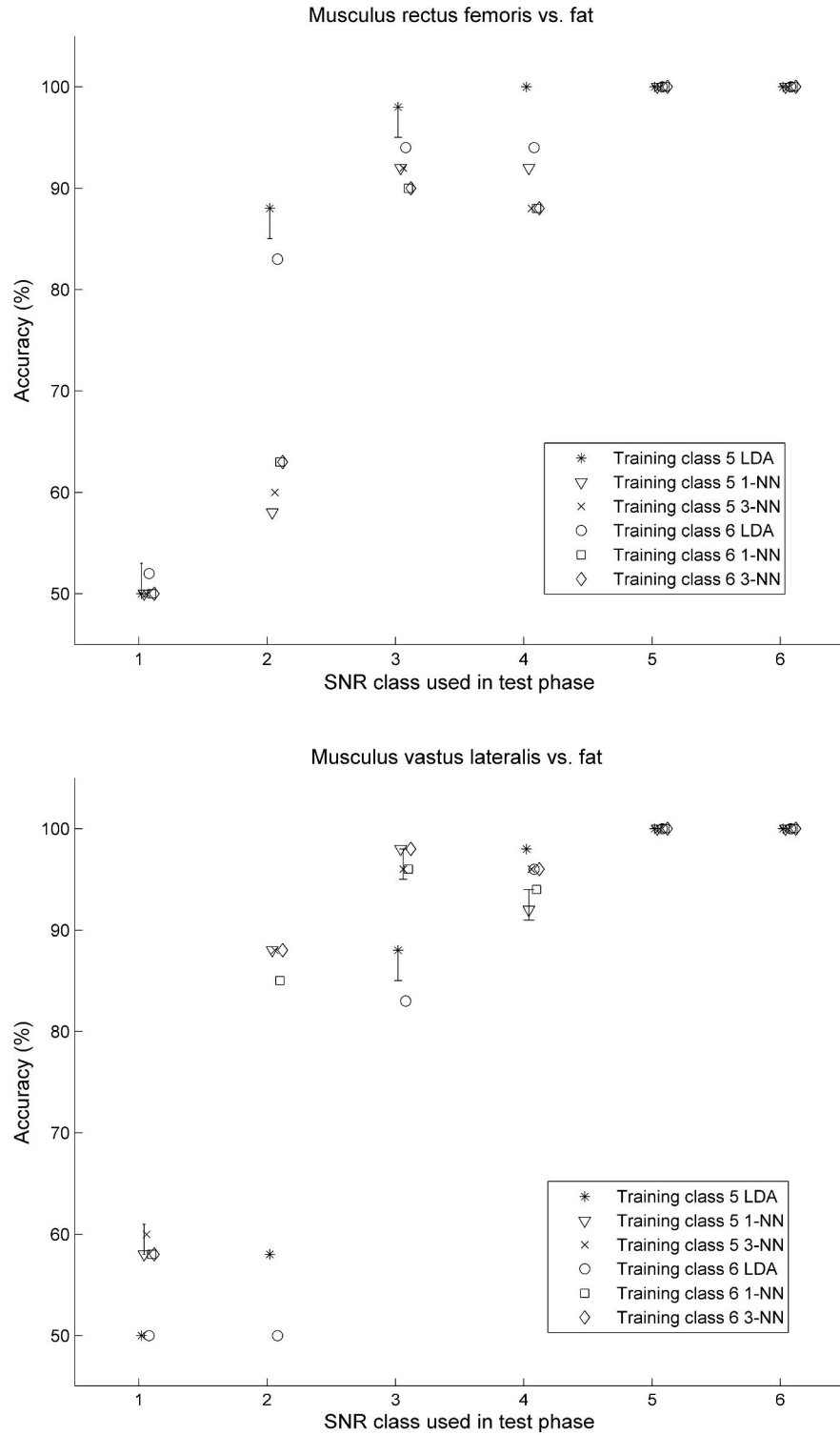


Figure 3. Results of LDA, 1-NN and 3-NN classifications of m. rectus femoris vs. fat (upper) and m. vastus lateralis vs. fat (lower). Tissues in SNR classes 5 (second highest SNR) and 6 (highest SNR) were used for training. Test classes from 1 (lowest SNR) to 6 (highest SNR) are shown on x-axis and test accuracies (%) on y-axis.

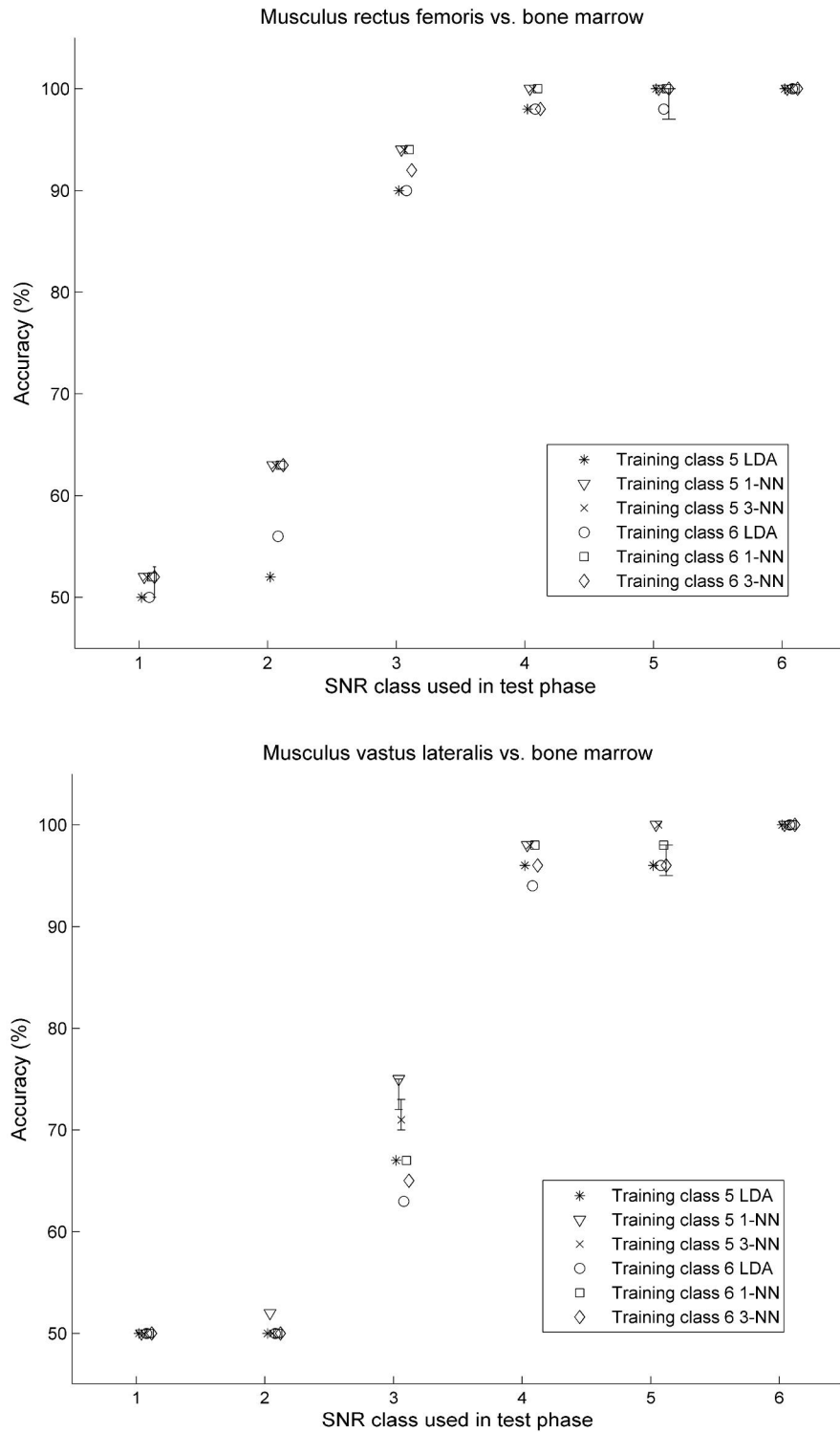


Figure 4. Results of LDA, 1-NN and 3-NN classifications of m. rectus femoris vs. bone marrow (upper) and m. vastus lateralis vs. bone marrow (lower). Tissues in SNR classes 5 (second highest SNR) and 6 (highest SNR) were used for training. Test classes from 1 (lowest SNR) to 6 (highest SNR) are shown on x-axis and test accuracies (%) on y-axis.

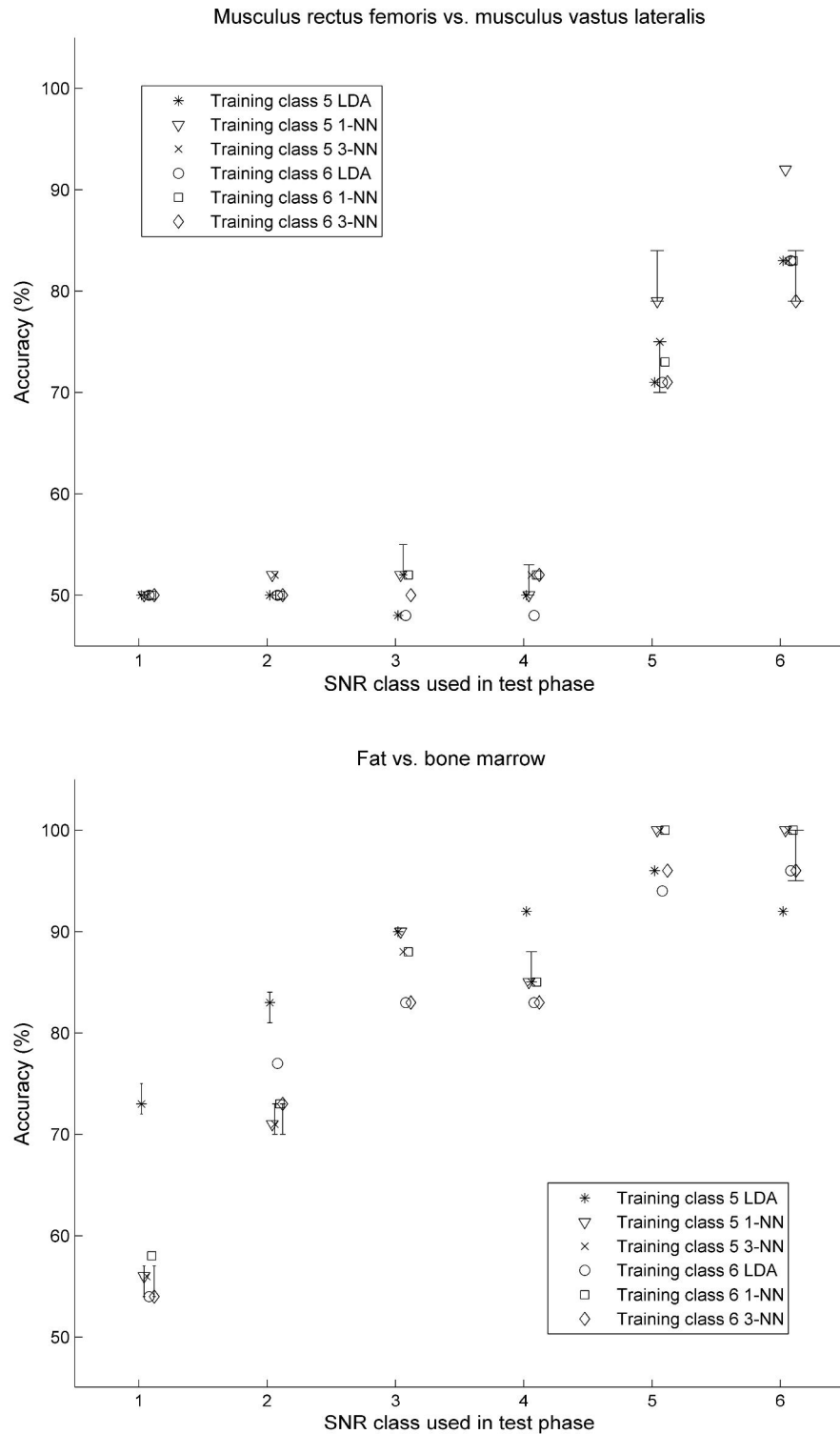


Figure 5. Results of LDA, 1-NN and 3-NN classifications of m. rectus femoris vs. m. vastus lateralis (upper) and fat vs. bone marrow (lower). Tissues in SNR classes 5 (second highest SNR) and 6 (highest SNR) were used for training. Test classes from 1 (lowest SNR) to 6 (highest SNR) are shown on x-axis and test accuracies (%) on y-axis.

4. DISCUSSION

We excluded histogram based parameters in this study, thus differing from the approach of Skoch et al. [10] This exclusion should make the classification results more reliable. The best classification results were obtained when fat was compared with m. rectus femoris and m. vastus lateralis in our test cases. (See Figure 3 for details.) When the SNR of ROIs included in the test classes was above 20 dB, the accuracy of all but a few classifications reached 90% or more. Moreover, when SNR was above 30 dB, practically no errors were made in the separation of fat and muscle tissues. Respectively, below 15 dB the results deteriorated quickly and fell close to 50% accuracy. This behaviour can be compared with [9], where the classification accuracy of different brain structures by means of texture analysis was almost perfect with high SNR but declined sharply when the lowest SNR classes were examined.

A clear difference between classification methods was not found. When m. rectus femoris and fat were compared with each other, LDA seemed to outweigh the nearest neighbour based methods, but in the separation of m. vastus lateralis and fat the nearest neighbour approach gave slightly better results.

As can be seen in Figure 4, bone marrow was almost as easily separable as fat tissue from m. rectus femoris and m. vastus lateralis. The classification accuracy that was over 90% with SNR over 25 dB with muscles and 20 dB with bone marrow started to decrease rapidly when SNR fell. The classification accuracy of each method was below 65% when SNR belonged to the two lowest classes. The decrease in the accuracy seemed to be steeper than when fat and muscle tissues were compared.

When LDA was used to separate muscle tissues either from fat or bone marrow, the classification accuracy was at least 94%, if SNR was over 30 dB. The results were close to similar with the nearest neighbour classifiers as well, but some dispersion from LDA results was found in a couple of cases. The results remained relatively good in most of the test cases, when SNR was at least 19 dB for muscle and fat and at least 15 dB for bone marrow. With the two lowest SNR groups, the classification results were almost random.

The separability of fat and bone marrow was relatively good, above 90% with the two highest SNR groups (SNR at least 33 dB for fat and at least 30 dB for bone marrow) as shown in Figure 5. The classification results remained above 80%, until fat tissue SNR decreased below 19 dB and bone marrow SNR below 15 dB. The classification accuracy declined smoothly compared to the other tissue classifications. Even with the second lowest SNR class it reached 70%. In addition, the classification accuracy of LDA was above 70% with all SNR classes if the second highest SNR class was used for training.

M. rectus femoris and m. vastus lateralis were the hardest ones to separate from each other with all the classification methods used. (See Figure 5.) With the highest SNR class (SNR at least 37 dB), the classification accuracy of 1-NN was over 90% and close to 80% with the other methods. However, the accuracy decreased rapidly as a function of additional noise. This can be compared to the results of Skoch et al. [10] that showed 80% agreement between the radiological classification and the texture analysis of healthy and diseased calf muscles.

The results for the second highest SNR class (SNR over 30 dB) of muscle classification showed less than 80% accuracy, and with more noise the classification was basically a random process. The use of the nearest neighbour classifier (1-NN) with the second highest SNR training class gave clearly the best results, other methods showing closely the same accuracy with each other. It seems to be an inevitable conclusion that the separation of muscles with texture analysis is a very difficult task even if the noise level is low. This means that the classifier as well as the parameters must be selected very carefully.

In all comparisons, the classification accuracy increased sharply as a function of SNR except in the case of fat and bone marrow, where the increase was clearly smoother. In most of the other cases there seemed to be a clear SNR threshold, below which the comparison was mostly random. The same effect has been found by Schad and Lundervold [9]. According to them, some textural properties appear after noise has dropped below a certain level. Most often this was obtained with the two lowest SNR classes, but in comparison of m. rectus femoris and m. vastus lateralis only the two highest SNR classes (SNR over 30 dB) could be used, and even in those cases the separability was relatively low.

There were no large differences between the results of the LDA and NN methods, even though the variation in a couple of low SNR cases was very high, because the task of texture analysis becomes much harder when additional noise is present. The variation of accuracy between methods remained mostly within 10% units. However, the variation of

accuracies was especially high when m. rectus femoris and m. vastus lateralis were compared to fat with a relatively low SNR. Unfortunately, none of the methods was superior with both of these muscle types. (See Figure 3 for details.)

We found that the classification results are relatively dispersed when fat is compared to other tissue types. It can also be seen that the results of texture analysis were slightly better in some cases if the training class was the second highest SNR class instead of the highest one. The same effect with absolute gradient based texture parameters has been detected earlier in a phantom study by Mayerhoefer et al. [13] They found this effect somewhat paradoxical, but we suggest an explanation to this phenomenon. With lower SNR levels finer details of tissue structures cannot be seen at all, and the texture analysis becomes inaccurate. However, if the training is completed with images including a moderate level of noise, tissue characteristics of a larger scale may be taken into account and thus higher levels of noise do not hide so many of the tissue properties under investigation.

5. CONCLUSIONS

It is relatively hard to define SNR levels below which texture analysis becomes too inaccurate, because the exact use of texture analysis depends on the application. However, one may conclude that if the SNR level is too low, the accuracy of texture analysis is inadequate. As SNR increases, there is usually a threshold above which a rapid increase in accuracy can be obtained. In our study cases, the improvement of texture analysis accuracy was mostly seen between 20 dB and 30 dB SNR levels, depending on the investigated tissue. When SNR remains above 30 dB, 90% classification accuracy can be reached in most cases.

ACKNOWLEDGEMENTS

The Tampere City Science Fund supported this research.


REFERENCES

- [1] Hajek, M., Dezortova, M., Materka, A. and Lerski, R. (editors), [Texture Analysis for Magnetic Resonance Imaging], Med4 Publishing, Prague, 1-234 (2006).
- [2] Savio, S. J., Harrison, L. C. V., Luukkaala, T., Heinonen, T., Dastidar, P., Soimakallio, S. and Eskola, H. J., "Effect of slice thickness on brain magnetic resonance image texture analysis," *Biomed Eng Online* 2010, 9:60 (2010).
- [3] Laun, F. B., Schad, L. R., Klein, J. and Stieltjes, B., "How background noise shifts eigenvectors and increases eigenvalues in DTI," *Magn Reson Mater Phy* 22(3), 151-158 (2009).
- [4] Farrell, J. A. D., Landman, B. A., Jones, C. K., Smith, S. A., Prince, J. L., van Zijl, P. C. M. and Mori, S., "Effects of signal-to-noise ratio on the accuracy and reproducibility of diffusion tensor imaging-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5T," *J Magn Reson Imaging* 26(3), 756-767 (2007).
- [5] Erdogmus, D., Larsson, E. G., Yan, R., Principe J. C. and Fitzsimmons, J. R., "Measuring the signal-to-noise ratio in magnetic resonance imaging: a caveat," *Signal Process* 84(6), 1035 –1040 (2004).
- [6] Materka, A., Strzelecki, M., Lerski, R. and Schad, L., "Feature evaluation of texture test objects for magnetic resonance imaging," *Texture Analysis in Machine Vision, Series in Machine Perception and Artificial Intelligence –Vol. 40*, edited by Pietikainen, M. K., World Scientific Publishing Co. Pte. Ltd., Singapore, 197-206 (2000).
- [7] Jirak, D., Dezortova, M. and Hajek, M., "Phantoms for texture analysis of MR images. Long-term and multi-center study," *Med Phys* 31(3), 616-622 (2004).
- [8] Quantitation of magnetic resonance image texture. Final report. COST action B11. Period: from May 1998 to May 2002. Technical Committee Medicine and Health.
- [9] Schad, L. R. and Lundervold, A., "Influence of resolution and signal to noise ratio on MR image texture," in Hajek, M., Dezortova, M., Materka, A. and Lerski, R. (editors), [Texture Analysis for Magnetic Resonance Imaging], Med4 Publishing, Prague, 129-149 (2006).

- [10] Skoch, A., Jirak, D., Vyhnánovská, P., Dezortová, M., Fendrych, P., Rolencová, E. and Hájek, M., "Classification of calf muscle MR images by texture analysis," *Magma* 16(6), 259-267 (2004).
- [11] Dietrich, O., Heiland, S. and Sartor, K., "Noise correction for the exact determination of apparent diffusion coefficients at low SNR," *Magn Reson Med*, 45(3), 448-453 (2001).
- [12] <http://www.eletel.p.lodz.pl/mazda/index.php?action=mazda>
- [13] Mayerhoefer, M.E., Szomolanyi, P., Jirak, D., Materka, A. and Trattnig, S., "Effects of MRI Acquisition Parameter Variations and Protocol Heterogeneity on the Results of Texture Analysis and Pattern Discrimination: An Application-oriented Study," *Med Phys* 36(4), 1236-1243 (2009).

Hemispheric asymmetry measured by texture analysis and diffusion tensor imaging in two multiple sclerosis subtypes

Sami Savio^{1,2}, Ullamari Hakulinen^{1,2}, Pertti Ryymin²,
Sanna Hagman³, Prasun Dastidar², Seppo Soimakallio²,
Irina Elovaara^{3,4} and Hannu Eskola^{1,2}

Acta Radiologica
2015, Vol. 56(7) 844–851
© The Foundation Acta Radiologica
2014
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0284185114539323
acr.sagepub.com


Abstract

Background: This paper addresses two subtypes of multiple sclerosis (MS), primary progressive multiple sclerosis (PPMS) and relapsing-remitting multiple sclerosis (RRMS). The separation of PPMS and RRMS is challenging in certain cases.

Purpose: To quantitatively determine MS subtypes using texture analysis (TA) and diffusion tensor imaging (DTI).

Material and Methods: T1-weighted (TIW) magnetic resonance imaging (MRI) and DTI of the left and right brain hemispheres of 17 patients with PPMS and 19 patients with RRMS were studied. Areas of the caudate nucleus and thalamus were investigated as normal appearing gray matter (NAGM), and areas of the cerebral peduncle and centrum semiovale were investigated as normal appearing white matter (NAWM). The described locations were symmetrical and were accurately marked. TA was performed on the TIW images, and the fractional anisotropy and apparent diffusion coefficient were determined from the DTI data.

Results: Hemispherical differences were found with both TA and DTI. Several texture and diffusion tensor parameter values calculated for the left and right hemispheres of the patients showed statistically significant differences. The patients with RRMS had greater significant differences ($P < 0.01$) in the thalamus between the hemispheres than did the patients with PPMS. The TA classification accuracy of the PPMS and RRMS subtypes was above 80%.

Conclusion: TA can be helpful when distinguishing between PPMS and RRMS, while DTI appears to reveal the hemispherical asymmetry of RRMS patients.

Keywords

Magnetic resonance imaging (MRI), MR diffusion, brain, tissue characterization

Date received: 19 February 2014; accepted: 13 May 2014

Introduction

Multiple sclerosis (MS) is a common autoimmune disease of the central nervous system characterized by several complex pathological processes, for example, inflammation and demyelination (1). The revised McDonald (2005) criteria (2) assess the value of cerebrospinal fluid analysis in diagnosing primary progressive multiple sclerosis (PPMS), which is characterized by a steady progression of irreversible disability. A more common subtype of the disease is relapsing-remitting multiple sclerosis (RRMS). Patients with RRMS suffer from episodes of acute exacerbation, followed by complete or partial recovery. However, the worsening symptoms in PPMS sometimes resemble a relapse,

so the separation of PPMS and RRMS may be challenging.

¹Department of Electronics and Communications Engineering, Tampere University of Technology, Tampere, Finland

²Department of Radiology, Tampere University Hospital, Tampere, Finland

³Neuroimmunology Unit, Medical School, University of Tampere, Tampere, Finland

⁴Department of Neurology, Tampere University Hospital, Tampere, Finland

Corresponding author:

Sami Savio, Department of Radiology, Tampere University Hospital, Teiskontie 35, FIN-33521 Tampere, Finland.
Email: sami.savio@tut.fi

According to Filippi (3), priority should be given to the analysis of normal appearing white matter (NAWM). Damage in normal appearing white and gray matter is an early feature of MS (4), and as a concept of clinical diagnosis, NAWM has been investigated in previous studies (3,5). Normal appearing gray matter (NAGM) has also been assessed using modern techniques (6,7), but not as widely as NAWM.

Texture analysis (TA) is a structural approach based on patterns in the image. It is a useful method for finding spatial features in medical images that cannot be revealed by human visual perception. The usability of TA has already been shown in a wide range of medical as well as other applications (8). However, the properties of the textures and patterns vary greatly between applications. The most common approach to TA uses statistical methods. Several different types of texture parameters can be calculated, including the histogram, co-occurrence, run-length, gradient, autoregressive, and wavelet-based parameters. The findings from previous MS studies (9,10) indicated that the plaques show a clear and distinguishable texture. Among others, Zhang et al. (9) and Harrison et al. (10) have used TA in the separation of MS plaques and NAWM. Although run-length matrix and wavelet-based methods may also be useful, the co-occurrence matrix method is likely the most commonly used texture analysis method in the investigation of the MS disease (11).

Diffusion tensor imaging (DTI) is a magnetic resonance imaging (MRI) method that tracks the diffusion of water molecules, thus indicating directional data in a three-dimensional space (12). DTI can be considered a functional method. The diffusion tensor matrix can be generated from multidirectional diffusion-weighted data. Fractional anisotropy (FA) measures diffusion directionality, whereas the apparent diffusion coefficient (ADC) refers to the diffusion capacity of the tissue. The applications of DTI vary from the evaluation of dementia (13) to the assessment of lesions in epilepsy (14).

According to Roosendaal et al. (15), DTI measures such as FA are sensitive to NAWM areas, including the corpus callosum and the corona radiata. In addition, Commowick et al. (5) reported that DTI measures enable the early detection of the extent of MS disease around the lesions. MS may also lead to hemispherical asymmetries in the normal appearing matter (16). In one of our earlier studies (17), small hemispherical asymmetries of the brain of healthy volunteers were also reported in regions such as the corona radiata and the posterior limb of the internal capsule, but the asymmetries were typically smaller in control individuals than in MS patients, as shown by Reich et al. (16).

It was hypothesized in this study that TA and DTI may complement each other in the classification of

different MS subtypes. For NAGM, the caudate nucleus and thalamus were selected as the studied locations, and for NAWM, the cerebral peduncle and centrum semiovale as NAWM were selected. The texture parameters and DTI (FA and ADC) values of PPMS and RRMS patients were evaluated in these locations. The measurements also included one MS plaque per patient. The aim of this study was to combine TA and DTI to enable an enhanced separation of PPMS and RRMS, thus improving the specificity of MS subtypes, the associated treatment, and the long-term prognosis of the patients.

Material and Methods

This study was conducted in the Imaging Centre of the Pirkanmaa Hospital District, Finland, and it was approved by the Regional Ethics Committee of the Expert Responsibility area of Tampere University Hospital (code R05157). The patients gave written informed consent for their participation.

Patients

This study included 17 patients with PPMS (age range, 42–72 years) and 19 patients with RRMS (age range, 22–53 years). All patients were neurologically examined with the help of the expanded disability status scale developed by Kurtzke (18), the number of relapses 2 years before the study, the disease duration, the progression index, and immunomodulatory treatments. The diagnosis of MS was based on the revised McDonald criteria (2005), which were developed by Polman et al. (2). All patients met the criteria (Table 1).

Table 1. Clinical characteristics of patients with primary progressive multiple sclerosis (PPMS) and relapsing-remitting multiple sclerosis (RRMS).

Clinical characteristic	PPMS	RRMS
Patients (n)	17	19
Age (years)*	58 (54–65)	38 (29–44)
Gender (F/M)	11/6	12/7
EDSS*	6.0 (4.0–6.0)	2.0 (1.0–2.4)
Time from the diagnosis (years)*	14.2 (9.7–17.7)	3.2 (0.7–6.1)
Time from the first symptom (years)*	16.5 (15.2–24.3)	8.2 (4.5–12.3)

*Median (25th–75th percentiles in parentheses).

Magnetic resonance imaging

All patients underwent MRI examination. The patient MRI images were acquired with a Siemens 1.5 T MR scanner (Siemens, Avanto SQ, Siemens Healthcare, Erlangen, Germany). Loevner et al. (19) have shown that T1-weighted (T1W) plaques represent axonal degeneration and are clearly associated with disease progress and invalidity in MS. The following magnetization-prepared rapid gradient echo T1W MR parameters were used: repetition time (TR), 1160 ms; echo time (TE), 4.24 ms; inversion time (TI), 600 ms; flip angle, 15 degrees; slice thickness, 0.9 mm; (axial) in-plane resolution, 0.45×0.45 mm; field of view (FOV), 230×200 mm, and acquisition matrix size, 512×448 pixels.

The DTI images were acquired during the same imaging session as the T1W images. Typical parameters for the DTI single-shot diffusion-weighted echo-planar imaging sequence were: TR, 3500 ms; TE, 96 ms; FOV, 230×230 mm; matrix size, 128×128 ; averages, 3; slice/gap, 5.0/1.5 mm; voxel dimension, $1.8 \times 1.8 \times 5.0$ mm; b-factors, 0 and 1000 s/mm^2 ; and 12 gradient orientations.

Image analysis

For evaluation of NAGM, slices in the anatomical levels of the caudate nucleus and thalamus were selected, as shown in Fig. 1. In addition, slices at the levels of the cerebral peduncle and centrum semiovale

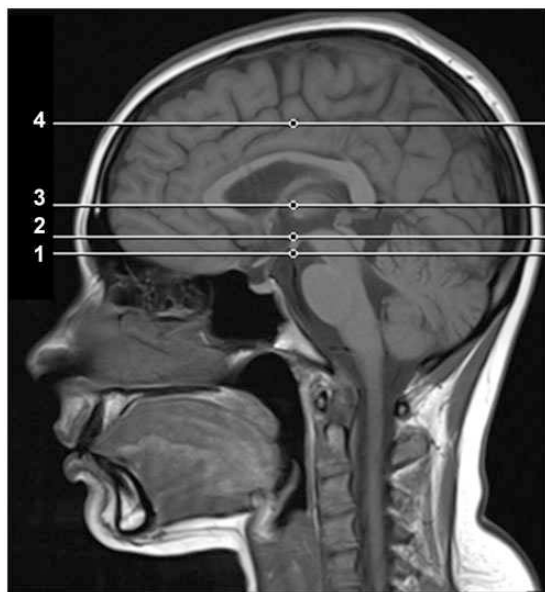


Fig. 1. An example of the axial T1W slice selection. The studied slice levels correspond to the [1] cerebral peduncle, [2] caudate nucleus, [3] thalamus, and [4] centrum semiovale. Similar locations were chosen for diffusion tensor imaging analysis.

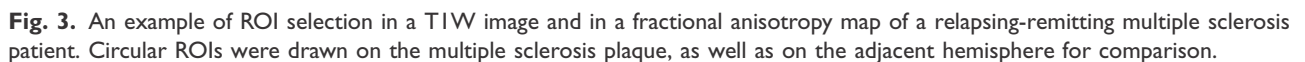
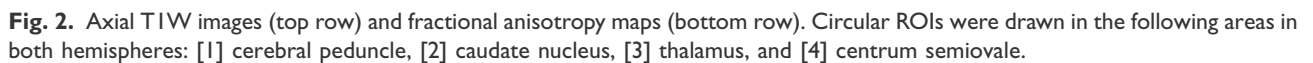
were selected to investigate NAWM. MRI and DTI slices were selected simultaneously on the Siemens Syngo Multimodality Workplace (Siemens Healthcare, Malvern, PA, USA). Circular regions of interest (ROIs) were manually placed on axial T1W images with the TA toolkit MaZda v. 3.3, software that was developed at the Technical University of Lodz, Lodz, Poland, for MRI TA in association with the COST B11 European Project “Quantitative Analysis of Magnetic Resonance Image Texture” (8), and on DTI images with the Siemens Syngo Multimodality Workplace Neuro3D Engine (Siemens Healthcare, Malvern, PA, USA) (Fig. 2).

The actual size of the ROI ($18\text{--}82 \text{ mm}^2$) depended on the anatomical location but was the same in the measurements of the anatomical MR data and the diffusion tensor data. ROIs with an actual size of 18 mm^2 were also drawn on hypointense MS lesions and in the adjacent hemisphere for comparative purposes. Only hypointense, irreversible lesions represented by black holes (permanent axonal degeneration) on T1W images were selected. An example of an image with ROIs drawn on an MS plaque and on a corresponding location in the other hemisphere is shown in Fig. 3.

The gray level intensity was normalized for each image within the range $(\mu - 3\sigma, \mu + 3\sigma)$, where μ is the mean gray level value, and σ is the standard deviation. After normalization, 132 texture parameters in each ROI were calculated in total; these parameters consisted of the co-occurrence matrix-based parameters of 11 subtypes with a direction of 0, 45, 90, or 135 degrees and a pixel distance of 1, 2, or 3 pixels. The size of the ROI was approximately 100 pixels in many cases; thus, the pixel distance was limited to 3 to improve statistical credibility. All texture parameters were calculated with MaZda. The FA and ADC values were given by the Siemens Syngo Multimodality Workplace Neuro3D Engine.

Statistical analysis

The texture parameter values for the left and right hemispheres were compared in the cerebral peduncle, caudate nucleus, thalamus, and centrum semiovale to reveal possible differences between the hemispheres of patients with the two MS subtypes. The studied dataset was relatively small, and a normal distribution of data could not be assumed; thus, this comparison was made using the non-parametric Wilcoxon signed-rank test. Additionally, *P* values based on the Wilcoxon signed-rank test were calculated for the DTI data. These values were determined for the aforementioned four anatomical regions and the FA and ADC parameters of both MS subtypes; the combinations provided 16 *P* values. To ensure an adequate level of statistical significance,



When selecting the most useful parameters for the separation of classes, one can utilize the Fisher coefficient to compare the between-classes variance to the within-class variance. The Fisher coefficient is a common and reliable parameter selection method and is the gold standard in many texture applications (8).

To improve the reliability of the results, k-fold cross validation (with $k=18$) was used when testing the

Table 2. Results of texture analysis. The proportion (%) of texture parameters with statistically significant differences ($P < 0.01$) between the left and right hemispheres. The P values for the 132 texture parameters were obtained using the Wilcoxon signed-rank test. Four brain areas in patients with the primary progressive multiple sclerosis (PPMS) and relapsing-remitting multiple sclerosis (RRMS) subtypes were examined.

Region	Proportion (%) of statistically significant parameters	
	PPMS	RRMS
Cerebral peduncle	1	1
Thalamus	1	23
Caudate nucleus	2	1
Centrum semiovale	0	2

algorithmic accuracy. This validation provides an estimate of the results that can be expected when the algorithm is tested with unseen data (20). In the cross-validation, 1000 classification rounds were run with separate training and test sets. Each set consisted of approximately one half of the PPMS and RRMS patients' ROIs. The median value of these classification results was considered a measure of the real power of the classification algorithm.

Results

TA and DTI analyses: differences between the left and right hemispheres of the brain

The numbers and proportions of texture parameters with a significance level of $P < 0.01$ in the comparison of hemispheres are shown in Table 2. The proportion of statistically significant parameters was especially large in the thalamus of patients with RRMS; thus, the proportion of texture parameters in different groups with statistically significant differences ($P < 0.01$) in the left and right hemispheres for the thalamus of these patients are shown in Table 3.

The mean values and standard deviations for FA and ADC are shown in Table 4. The P values for the hemispherical differences in terms of the DTI data are shown in Table 5.

Correlation between the TA and DTI measurements

The lowest absolute value of correlation between the texture and DTI parameters was as low as 0.0005. This was the correlation between the texture parameter sum of entropy with 3-pixel distance and 90-degree direction with the ADC values of patients with PPMS.

The highest absolute single value of correlation (0.64) existed between the texture parameters and the

Table 3. Results of texture analysis. Proportion of texture parameters in different groups with statistically significant differences ($P < 0.01$) between the left and right hemispheres for the thalamus of patients with RRMS.

Texture parameter subtype	Proportion (%)
Sum of variance	58
Inverse difference moment	42
Correlation	33
Sum of entropy	33
Difference entropy	25
Contrast	25
Difference variance	25
Sum of squares	17
Sum average	0
Entropy	0
Angular second moment	0

ADC values for the patients with PPMS. A statistically significant correlation ($P < 0.01$) was not found between texture parameters and the DTI parameters.

The absolute values of the median correlations between the texture parameters and the DTI values, grouped by texture parameter subtypes, are shown in Table 6. No statistically significant correlations ($P < 0.01$) were found.

Classification accuracy between MS subtypes using TA and DTI

The classification results for the PPMS and RRMS subtypes with four classification methods (LDA, NDA, 1-NN, and 3-NN) and the two texture parameters with the highest Fisher coefficients (S(0,2) Entropy and S(2,2) SumEntropy) were calculated on the hypointense lesions (Table 7). In addition, the FA and ADC values were used in some classifications.

In an additional classification of the two MS subtypes, only the information measured in both hemispheres in the thalamus was included while the information regarding the patients' MS plaques was excluded. The results showed that, with two texture parameters, the classification accuracy increased to 82% with LDA.

Discussion

In this novel study, the potential differences between the data obtained for patients with PPMS and RRMS were examined and measured with TA and DTI methods. Our findings suggest that the TA and DTI

Table 4. DTI results. Mean values and standard deviations of diffusion tensor parameters in the cerebral peduncle, thalamus, caudate nucleus, and centrum semiovale.

Region	Parameter	PPMS		RRMS	
		Left hemisphere	Right hemisphere	Left hemisphere	Right hemisphere
Cerebral peduncle	FA	0.701 ± 0.056	0.688 ± 0.068	0.713 ± 0.052	0.704 ± 0.064
	ADC	0.803 ± 0.044	0.810 ± 0.052	0.794 ± 0.060	0.793 ± 0.063
Thalamus	FA	0.279 ± 0.075	0.309 ± 0.035	0.310 ± 0.039	0.305 ± 0.034
	ADC	0.797 ± 0.050	0.792 ± 0.050	0.754 ± 0.042	0.760 ± 0.039
Caudate nucleus	FA	0.216 ± 0.048	0.245 ± 0.055	0.221 ± 0.060	0.223 ± 0.043
	ADC	0.699 ± 0.060	0.722 ± 0.063	0.664 ± 0.045	0.656 ± 0.041
Centrum semiovale	FA	0.549 ± 0.072	0.539 ± 0.098	0.572 ± 0.098	0.571 ± 0.115
	ADC	0.704 ± 0.042	0.734 ± 0.053	0.724 ± 0.075	0.742 ± 0.083

Table 5. DTI results. *P* values acquired with the Wilcoxon signed-rank test showing differences between the left and right hemispheres of the brain in terms of diffusion tensor data.

Region	PPMS (<i>n</i> = 17)		RRMS (<i>n</i> = 19)	
	FA	ADC	FA	ADC
Cerebral peduncle	0.44	0.72	0.38	0.10
Thalamus	0.75	0.30	0.013*	0.39
Caudate nucleus	0.92	0.49	0.006 [†]	0.39
Centrum semiovale	0.94	0.14	0.87	0.025*

*Statistical significance *P* < 0.05.

[†]Statistical significance *P* < 0.01.

ADC, apparent diffusion coefficient (10⁻³ mm²/s); FA, fractional anisotropy; PPMS, primary progressive multiple sclerosis; RRMS, relapsing-remitting multiple sclerosis.

methods can detect a difference between the PPMS and RRMS subtypes.

The left and right hemispherical texture in the thalamus was different in the RRMS patient group, while texture differences between the hemispheres were not statistically significant in the patients with PPMS. In the DTI analysis, statistically significant differences between the left and right hemispheres were also found in patients with the RRMS subtype, especially in the caudate nucleus, in addition to the observation of a hemispherical difference in the centrum semiovale and the thalamus. The hemispherical differences were greater in RRMS than in PPMS. These differences existed in both the NAGM and NAWM areas, but the statistical significance was higher in the NAGM.

According to several earlier studies (21,22), cerebral degeneration occurs in the deep gray matter and may even precede the development of the white matter damage (23). For example, Zhou et al. (21) evaluated FA values in hyperintense MS plaques in the deep gray

Table 6. Correlation of TA and DTI. Absolute values of the median correlation between the texture parameters and the fractional anisotropy (FA) and apparent diffusion coefficient (ADC) values in primary progressive multiple sclerosis (PPMS) and relapsing-remitting multiple sclerosis (RRMS).

Texture parameter subtype	PPMS (<i>n</i> = 17)		RRMS (<i>n</i> = 19)	
	FA	ADC	FA	ADC
Sum of variance	0.026	0.14	0.14	0.32
Inverse difference moment	0.30	0.19	0.055	0.049
Correlation	0.12	0.028	0.15	0.20
Sum of entropy	0.29	0.27	0.27	0.24
Difference entropy	0.049	0.023	0.089	0.044
Contrast	0.21	0.049	0.10	0.012
Difference variance	0.15	0.13	0.14	0.14
Sum of squares	0.25	0.28	0.15	0.42
Sum average	0.010	0.29	0.36	0.38
Entropy	0.37	0.35	0.19	0.19
Angular second moment	0.45	0.32	0.18	0.16

matter and found that hyperintense lesions had higher FA values than those of the hypointense plaques, but lower than those of the NAWM and normal white matter in the controls. Calabrese et al. (22) showed that increased FA in gray matter may indicate microglial activation or other inflammatory events.

Riccitelli et al. (24) used T2-weighted MR images to investigate MS subtypes and found that cognitively impaired patients with RRMS had significant gray matter loss in the left thalamus relative to that in patients with PPMS. They suggested that different pathological mechanisms explain the difference between the PPMS and RRMS subtypes of MS. Fabiano et al. (25) found that the ADC values in the left and right thalamus differ in patients suffering from RRMS.

Table 7. Classification accuracy of selected TA and DTI parameters for primary progressive multiple sclerosis (PPMS) and relapsing-remitting multiple sclerosis (RRMS) with linear discriminant analysis (LDA), non-linear discriminant analysis (NDA), 1-nearest neighbor (1-NN), and 3-nearest neighbors (3-NN). The classification was based on the two texture parameters with the highest Fisher coefficients (left column) and on the combination of the same texture parameters with two DTI parameters (fractional anisotropy [FA] and the apparent diffusion coefficient [ADC] values, right column).

Algorithm	Accuracy (%) (2 texture parameters)	Accuracy (%) (2 texture parameters + FA + ADC)
LDA	73	77
NDA	81	79
1-NN	80	81
3-NN	81	79

In a recent study, Minagar et al. (26) recommended that particular attention be paid to the changes in thalamus, as it is involved in a wide range of neurological functions. Hasan et al. (27) stated that although there were no visible thalamic MS lesions, the loss of thalamic macrostructural volume and microstructural integrity may be related to connections to the cortex and other brain regions through white matter pathways and can be detected due to the increased diffusivity and increased T2 relaxation time. Graulieres et al. (28) reported that in the gray matter area of patients with RRMS, the shape of the ADC and FA value histograms may show a significant shift toward abnormal values, although the clinical progression might not be measurable. Their finding can be compared to our study, in which the measured FA values in the thalamus area of patients with RRMS showed statistically significant hemispherical differences (Tables 4 and 5); however, the same effect was not observed for the ADC values.

Our study showed a classification accuracy of 82%. Contrary to our hypothesis, the simultaneous application of TA and DTI did not improve the classification. The nearest neighbor classifiers provided the best results, followed closely by NDA and LDA. The use of the FA and ADC values in addition to the texture parameters did not significantly improve the classification results, although statistically significant correlations ($P < 0.01$) were not found between the texture parameter groups and the FA or ADC values.

Harrison et al. (10) stated that a clinical method for diagnosing MS in terms of TA may be developed in the future; this viewpoint was recently noted by Zhang (11). In addition, Zhang et al. (29) summarized the application studies of TA in MS by concluding that TA shows

promise for the extraction of clinically meaningful information from routine conventional MRI and for the characterization of disease activity and progression.

TA and DTI can be used as well in other degenerative disorders, for example, Parkinson's disease, cerebrovascular disease, and dementia. It would be interesting to investigate whether TA and DTI parameters can be further combined with any other type of image analysis method in the separation of MS subtypes, including RRMS and PPMS, and thus help clinicians in prescribing a suitable treatment to their patients soon after the first symptoms are observed.

This study does have limitations that are worth mentioning. There may be some variation in the parameter values for several reasons, including slice thickness and ROI placement. The slice thickness of the images ranged from 0.9 mm in the T1W images to 5.0 mm in the diffusion tensor data in our study. Although a partial volume effect may have affected some differences in the parameter values, in a previous study we showed that 1 mm and 3 mm thick slices are interchangeable in MS TA (30). Hakulinen et al. (17) studied the effect of intra- and inter-observer variability and repeatability of DTI ROI selection, and they reported that the repeatability was either excellent or appropriate in the majority of the cases they examined. Finally, Harrison et al. (10) reported that TA is an appropriate tool in terms of robustness for diagnosing MS patients with different algorithms and assumptions, including texture in consequent slices. The aforementioned findings support our assumption that neither the difference in slice thickness nor the possible variation in the ROI placement has a major effect on the results of our study.

In conclusion, several texture parameters and diffusion tensor parameter values in both brain hemispheres of patients with MS were statistically significantly different, enabling classification of MS subtypes. TA can be helpful when distinguishing between PPMS and RRMS, while DTI appears to reveal the hemispherical asymmetry in RRMS patients.

Conflict of interest

None declared.

Funding

This work was supported by research grants from The Finnish Cultural Foundation (grant number 00110850) and The South Savo Regional Fund (grant number 12121328).

References

- McQualter JL, Bernard CC. Multiple sclerosis: a battle between destruction and repair. *J Neurochem* 2007;100: 295–306.

2. Polman CH, Reingold SC, Edan G, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald criteria". *Ann Neurol* 2005;58:840–846.
3. Filippi M. In-vivo tissue characterization of multiple sclerosis and other white matter diseases using magnetic resonance based techniques. *J Neurol* 2001;248:1019–1029.
4. Inglese M, Salvi F, Iannucci G, et al. Magnetization transfer and diffusion tensor MR imaging of acute disseminated encephalomyelitis. *Am J Neuroradiol* 2002;23:267–272.
5. Commowick O, Fillard P, Clatz O, et al. Detection of DTI white matter abnormalities in multiple sclerosis patients. *Med Image Comput Comput Assist Interv* 2008;11:975–982.
6. Griffin CM, Chard DT, Ciccarelli O, et al. Diffusion tensor imaging in early relapsing-remitting multiple sclerosis. *Mult Scler* 2001;7:290–297.
7. Vrenken H, Pouwels PJW, Geurts JJG, et al. Altered diffusion tensor in multiple sclerosis normal-appearing brain tissue: cortical diffusion changes seem related to clinical deterioration. *J Magn Reson Imaging* 2006;23:628–636.
8. Hajek M, Dezortova M, Materka A, et al (eds) *Texture analysis for magnetic resonance imaging*. Prague: Med4Publishing, 2006.
9. Zhang Y, Zhu H, Ferrari R, et al. Texture analysis of MR images of minocycline treated MS patients. *Lect Notes Comp Sci* 2003;2878:786–793.
10. Harrison LCV, Raunio M, Holli KK, et al. MRI texture analysis in multiple sclerosis: Toward a clinical analysis protocol. *Acad Radiol* 2010;17:696–707.
11. Zhang Y. MRI texture analysis in multiple sclerosis. *Int J Biomed Imaging* 2012;2012:2.
12. Filler A. Magnetic resonance neurography and diffusion tensor imaging: origins, history, and clinical impact of the first 50,000 cases with an assessment of efficacy and utility in a prospective 5,000 patient study group. *Neurosurgery* 2009;65:A29–A43.
13. Chua TC, Wen W, Slavin MJ, et al. Diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease: a review. *Curr Opin Neurol* 2008;21:83–92.
14. Diehl B, Busch RM, Duncan JS, et al. Abnormalities in diffusion tensor imaging of the uncinate fasciculus relate to reduced memory in temporal lobe epilepsy. *Epilepsia* 2008;49:1409–1418.
15. Roosendaal SD, Geurts JJ, Vrenken H, et al. Regional DTI differences in multiple sclerosis patients. *Neuroimage* 2009;44:1397–1403.
16. Reich DS, Smith SA, Zackowski KM, et al. Multiparametric magnetic resonance imaging analysis of the corticospinal tract in multiple sclerosis. *Neuroimage* 2007;38:271–279.
17. Hakulinen U, Brander A, Ryymin P, et al. Repeatability and variation of region-of-interest methods using quantitative diffusion tensor MR imaging of the brain. *BMC Med Imaging* 2012;12:30.
18. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983;33:1444–1452.
19. Loevner LA, Grossman RI, McGowan JC, et al. Characterization of multiple sclerosis plaques with T1-weighted MR and quantitative magnetization transfer. *Am J Neuroradiol* 1995;6:1473–1479.
20. Kohavi R, John GH. Automatic parameter selection by minimizing estimated error. In: *Proceedings of the Twelfth International Conference on Machine Learning* 1995:304–312.
21. Zhou F, Shiroishi M, Gong H, et al. Multiple sclerosis: Hyperintense lesions in the brain on T1-weighted MR images assessed by diffusion tensor imaging. *J Magn Reson Imaging* 2010;31:789–795.
22. Calabrese M, Rinaldi F, Seppi D, et al. Cortical diffusion-tensor imaging abnormalities in multiple sclerosis: a 3-year longitudinal study. *Radiology* 2011;261:891–898.
23. Chard D, Miller D. Grey matter pathology in clinically early multiple sclerosis: evidence from magnetic resonance imaging. *J Neurol Sci* 2009;282:5–11.
24. Riccitelli G, Rocca MA, Pagani E, et al. Cognitive impairment in multiple sclerosis is associated to different patterns of gray matter atrophy according to clinical phenotype. *Hum Brain Mapp* 2011;32:1535–1543.
25. Fabiano AJ, Sharma J, Weinstock-Guttman B, et al. Thalamic involvement in multiple sclerosis: a diffusion-weighted magnetic resonance imaging study. *J Neuroimaging* 2003;13:307–314.
26. Minagar A, Barnett MH, Benedict RH, et al. The thalamus and multiple sclerosis: modern view of pathologic, imaging, and clinical aspects. *Neurology* 2013;80:210–219.
27. Hasan KH, Walimuni IS, Abid H, et al. Multimodal quantitative magnetic resonance imaging of thalamic development and aging across the human lifespan: implications to neurodegeneration in multiple sclerosis. *J Neurosci* 2011;31:16826–16832.
28. Graulieres E, Lotterie JA, Cassol E, et al. Relevance of the skewness index in DTI exploration of multiple sclerosis. *Magn Reson Mater Phy* 2009;22:89–100.
29. Zhang J, Tong L, Wang L, et al. Texture analysis of multiple sclerosis: a comparative study. *Magn Reson Imaging* 2008;26:1160–1166.
30. Savio SJ, Harrison LCV, Luukkaala, et al. Effect of slice thickness on brain magnetic resonance image texture analysis. *Biomed Eng Online* 2010;9:60.